



Freie Universität zu Berlin

Fachbereich Erziehungswissenschaften und Psychologie

SYSTEMIC PERSPECTIVES ON EMPATHY:
AN EMPIRICAL INVESTIGATION ON THE INFLUENCE OF GROUP
PARAMETERS ON EMPATHIC PROCESSES IN SMALL GROUPS

Dissertation

Zur Erlangung des akademischen Grades Doktor der Philosophie (Dr. phil.)

vorgelegt von:

Dipl.-Psych. Peter Paul Zurek

Betreuer und Gutachter:

Prof. Dr. Herbert Scheithauer

Prof. Dr. Christian von Scheve

Einreichung: 21. Mai 2015

Disputation: 12. Oktober 2015

Table of Contents

Acknowledgements.....	v
1. Introduction.....	1
2. Theoretical Considerations	4
2.1 Conceptualization of Empathy	4
2.1.1 Definition.....	5
2.1.2 The Multidimensional and Multifaceted Structure of Empathy	8
2.1.3 Functions of Empathy	14
2.2 Distinguishing Empathy from Related Constructs.....	16
2.2.1 Empathy and Theory of Mind.....	16
2.2.2 Empathy and Emotional Contagion.....	20
2.2.3 Empathy and Compassion.....	23
2.2.4 Empathy and Emotional Mimicry	24
2.2.5 Empathy and Perspective Taking	26
2.3 Trajectories of Empathy	27
2.3.1 Phylogenesis.....	27
2.3.2 Ontogenesis	30
2.3.3 Neurobiology.....	34
2.3.4 Social Learning	36
2.4 Empathy in the Social Context	38
2.4.1 Differentiation Between Individual and Relational Traits	38
2.4.2 Empathy as a Dyadic and Systemic Trait	43
2.4.3 The Role of Empathy in Social Groups	46
2.4.4 Interrelationship Between Group Parameters and Empathy.....	48
2.4.4.1 Diversity	48
2.4.4.1 Shared Goals.....	50
3. Central Research Questions and Hypotheses.....	52
4. Method	56
4.1 Participants	56
4.1.1 Sample Size and Composition	56

4.1.2 Demographic Characteristics	58
4.2 Procedure	60
4.2.1 First Experimental Session	60
4.2.2 Second Experimental Session.....	62
4.2.3 Third Experimental Session	64
4.2.4 Fourth Experimental Session.....	67
4.3 Measures.....	68
4.3.1 Questionnaires	68
4.3.1.1 IRI.....	68
4.3.1.2 CEEQ-GROUP	69
4.3.1.3 SDS.....	70
4.3.1.4 NEO-FFI.....	70
4.3.1.5 Questionnaire on Group Processes	71
4.3.1.6 Questionnaire on Group Composition.....	71
4.3.1.7 Questionnaire on Goal Representation	72
4.3.1.8 Questionnaire on Task Motivation (Shared Goal Condition).....	73
4.3.2 Multifaceted Empathy Test.....	73
4.3.3 Public Goods Game	74
4.3.4 Group Discussion Parameters	75
4.4 Statistical Analyses	78
4.4.1 Data Preparation	79
4.4.2 Preliminary Analyses	81
4.4.3 Testing for the Hypotheses	84
4.4.4 Follow-Up Analyses	85
5. Results	86
5.1 Descriptive Statistics and Reliabilities	86
5.1.1 Empathy Measures.....	86
5.1.2 Control Measures.....	91
5.2 Manipulation Check	95
5.2.1 Perceived Surface-Level Diversity and Diversity Beliefs.....	95
5.2.2 Goal Representation and Goal Commitment.....	99
5.2.3 Task Motivation and Orientation	102

5.3 Associations Between Empathy Measures and Control Variables	103
5.3.1 Social Desirability	104
5.3.2 Big Five Personality Traits	105
5.3.3 Group Climate and Identification	107
5.3.4 Dominance and Affiliation.....	109
5.3.5 Perceived Diversity, Diversity Beliefs, and Goal Representation	110
5.3.6 Age, Gender, and Education.....	111
5.4 Associations Between Individual and Systemic Empathy Measures.....	113
5.5 Impact of Group Parameters on Empathy	119
5.5.1 Questionnaires	119
5.5.2 Multifaceted Empathy Test	125
5.5.3 Public Goods Game	126
5.5.4 Group Discussion Parameters	128
5.6 Follow-Up Analyses on Possible Moderator and Mediator Effects	129
6. Discussion	137
6.1 Psychometric Quality of the Applied Measures	138
6.1.1 Objectivity	138
6.1.2 Internal Consistency.....	138
6.1.3 Stability	139
6.1.4 Construct Validity.....	142
6.2 Effectivity of the Experimental Manipulation	144
6.3 Longitudinal Effects of Group Parameters on Empathy	147
6.3.1 The Impact of Diversity on Empathy	147
6.3.2 The Impact of Shared Goals on Empathy	153
6.4 Correlates of Empathy.....	159
6.4.1 Individual States and Traits	159
6.4.2 Systemic States and Traits.....	169
6.5 On the Possibility of Antisocial Functions of Cognitive Empathy	171
6.6 Practical Implications	173
6.7 Limitations.....	177
6.8 Directions for Future Research	180
6.9 Main Conclusions and Final Remark	181

7. References	183
8. Appendix	205
8.1 Operationalization of Empathic Communication (Wacker, 2011)	205
8.2 Detailed Associations Between Empathy and Control Variables	206
8.3 Correlation Tables for Individual and Systemic Empathy Measures	221
8.4 Experimenter Manual	226
8.5 Debriefing	242
 List of Figures	 244
List of Tables	246
List of Tables in Appendix.....	248

Acknowledgements

First, I would like to express special gratitude to my mentor, Prof. Dr. Herbert Scheithauer, for his continuous support and excellent professional advice. I perceived his declaration to adopt this work and me in January 2013 as an empathic and truly altruistic act, as it is uncommon in everyday life and in the scientific community. From this moment on, his extraordinary competence, his tremendous psychological knowledge, the precision of his intellect, and also his unmatched skill in mentoring and motivating a student, have been inspiring, stimulating, and encouraging me throughout the course of this dissertation.

Further, I want to thank the members of the “Understanding Interaffectivity”-research group at the Cluster “Languages of Emotion”, especially Prof. Dr. Isabel Dziobek for mentoring this work in its first years, her patient support and input. I am also very grateful to my second advisor Prof. Dr. Christian von Scheve for his competent advice and his help to realize the interdisciplinary approach of this dissertation by sharing his great and profound sociological knowledge with me.

I also offer my thanks and appreciation to the management and staff of the Cluster “Languages of Emotion” at the Free University Berlin, for their organizational and technical support – especially Regula von Schintling, Anja Firsching, Dr. Markus Edler, and Stefan Liesch.

This dissertation was also much improved by discussions with colleagues, statisticians, and visiting scholars, especially Dipl.-Psych. Martin Schultze, Prof. Dr. Stefan Stürmer, Dr. Birte Siem, and Dipl.-Psych. Georg Hosoya, who should not remain unmentioned. I also owe my gratitude to the student assistants of our research group, especially Dipl.-Psych. Julia Bartz and Dipl.-Psych. Julia Großmann, not only for their help in conducting this study but also their inspiring input and feedback. Thanks must also go to Dipl.-Psych. Renata Wacker for her steady support and for sharing her profound knowledge on social psychology with me.

Moreover, I would like to express special gratitude to my family and my friends for steady support and encouragement in times of trouble. I would have never come so far without you all.

Finally and foremost, I thank my wife Florentine for having supported me emotionally and organizationally at any time. You made me experience firsthand the subject of my study: Without your empathy and love, this work would not have been possible.

1. Introduction

After World War II with all its horrors and over 60 million kills the human kind was confronted with fundamental questions, which challenge the very concept of humanity more strongly than ever before. How is it possible that human beings display such astonishing levels of destructiveness? How can we explain that human beings are capable of being merciless and cruel to this horrifying extent? And more importantly, can we do anything to prevent such destructive behaviors in the future? Finding definite answers to these questions seemed and seems to be an impossible task.

However, during the Nuremberg tribunals in 1946 the army psychologist Dr. Gustav Gilbert was entrusted with this demanding assignment. The U.S. Army asked him to study the minds and motivations of the leading Nazi defendants at the tribunals in order to find out what caused their behavior and decision making during the war. The results of his analyses are summarized in his book *"Nuremberg Diary"* (Gilbert, 1961). This diary was also the book on which the movie Nuremberg (directed by Yves Simoneau, 2000) was based.

In one scene of the movie, Dr. Gilbert (portrayed by Matt Craven) is telling the head prosecutor of the tribunals, Robert Jackson (portrayed by Alec Baldwin), the following:

"I told you once that I was searching for the nature of evil. I think I've come close to defining it. A lack of empathy. It's one characteristic that connects all the defendants. A genuine incapacity to feel with their fellow man. Evil, I think, is the absence of empathy."

Most of this monologue is a direct citation from the Nuremberg Diary (Gilbert, 1961). Even if considerations regarding the nature of evil are rather of philosophical nature, some aspects of Dr. Gilbert's conclusion still target the traditional domain of psychology. The primary goals of psychology as a scientific discipline are to describe, to explain, to predict, and to influence human behavior and experience (Wood, Wood, & Boyd, 2010). Therefore, it is the task of psychologists to provide answers on the following questions: Can antisocial behavior really be explained by a lack of empathy? Is it possible to predict such behavior by the level of empathic abilities and empathic concern? In reverse, can we as psychologists enhance empathy to prevent antisocial behavior? And, most importantly, what can be exactly understood by the term "empathy"?

There are no simple answers to these partly historical questions. However, the knowledge on the role of empathy in groups and societies can be deepened through theoretical and empirical investigations. Thus, this dissertation aims at providing a small contribution to the understanding of the broad and important concept of empathy, particularly in the context of social groups.

In the last century, emotional psychology and specifically the concept of empathy occupied an important position in commonsense psychology. Not only extreme human behaviors but also many other phenomena from our everyday lives are being explained by either a lack of empathy or by a strong empathic ability, translating into antisocial as well as prosocial behaviors. This is one of many reasons why the investigation of empathy has gained interest not only in psychological research, but also among researchers from several other disciplines, such as sociology, philosophy, neuroscience, history, ethology, and others (Preston & de Waal, 2002). However, each of these disciplines provides only a limited access towards a holistic understanding of empathy in the social context due to disciplinary boundaries. For example, the traditional psychological viewpoint on empathy is an individual psychological one: Empathy is regarded as a state or trait of one person in relation to another person. Contrary, sociology views empathy as a characteristic of groups or societies.

The present dissertation tries to provide an interdisciplinary perspective on empathy by combining perspectives from the domains of psychology and sociology. Specifically, it follows an operationalization of empathy as a characteristic of an individual *and* as a characteristic of a group or social system.

The following investigation will address the research question whether specific characteristics or parameters of a social group like diversity or mutual interdependence influence empathy on an individual and on a group level. Further, it will attempt to provide answers onto the following related research questions:

How can psychologists measure empathy on both an individual and on a systemic level? Is there a difference between individual and systemic empathy? What are the functional characteristics and effects of empathy in social groups? Is there a possibility to

influence empathic processes in these groups through psychological interventions? Are there also potential negative outcomes of empathy?

The present empirical investigation addresses these questions within a small group research design, which is based on both psychological and sociological theoretical background and a corresponding methodological approach derived from this interdisciplinary perspective on empathy.

In a first step towards realizing the research objectives, the following section provides theoretical background on the concept of empathy, its functions, related constructs, its trajectories, and its role in social groups. After deriving the hypotheses from these theoretical considerations, the small group design of the current study is presented in Section 4. Section 5 deals with the results of the investigation with regard to the multiple empathy measures that have been applied. The dissertation concludes with a critical discussion of the results, reflecting on shortcomings, limitations, and the practical relevance of this investigation, as well as deducting directions for future research on empathy in groups.

2. Theoretical Considerations

The following presentation of theoretical considerations is divided into five parts. First, definitions of the term empathy are discussed, as well as the multiple dimensions and facets of this construct and the psychological functions of empathy. In the second part, empathy is distinguished from related concepts such as theory of mind, emotional contagion, compassion, emotional mimicry, and perspective taking. The third section provides theoretical background on the phylogenetic, ontogenetic, neurobiological, and social determinants of empathy. Fourth, empathy is being discussed in the social context, as a systemic variable and with regard to the social functions of empathy. Finally, the interrelationship between empathy and specific group parameters is evaluated. The theoretical considerations conclude with a discussion of the central research question and hypotheses of this investigation.

2.1 Conceptualization of Empathy

As stated in the introduction, first of all it is important to make clear what empathy exactly means. The following chapter provides historical background on the concept of empathy followed by an elaborated discussion of how we can define and conceptualize this construct.

The word “empathy” is derived from the ancient Greek word ἐμπάθεια (empathia), which can be translated as strong affection or passion. However, the modern term “*empathy*” was first used by Edward Titchener (1867-1929) as an attempt to translate the German word “*Einfühlung*” (Stueber, 2008). The concept of “*Einfühlung*” was an important category in philosophical aesthetics of the 19th century and mainly explored by the German philosopher Theodor Lipps (1851-1914). Lipps understood “*Einfühlung*”, which means “*feeling into*” something in literal terms, as the ability to “*project yourself into what you observe*”. This projection could refer to a work of art, a piece of music, but also to another person: according to Lipps, the mind is able to mirror mental activities and experiences of another person, based on the observation of this person (Stueber, 2008).

Further, Lipps supposed that basic motor mimicry may be an underlying mechanism of conscious empathy (Hatfield, Cacioppo, & Rapson, 1993), thereby providing the theoretical background for the differentiation and functional interdependence between empathy and emotional contagion, which is still being discussed in current psychological research (see also Section 2.2.2).

However, it took almost one century after the theoretical works of Theodor Lipps and Edward Titchener until empirical evidence for this specific ability to mirror mental activities and experiences of another person was found. In 1992, the Italian neurophysiologist Giacomo Rizzolatti discovered in an experiment with monkeys that particular neurons are activated when a specific action is pursued, but also when the same specific action is being observed. Following experiments provided evidence for the same neural structures in humans. He called these neurons “mirror neurons” (see Keysers & Perrett, 2004; Rizzolatti & Craighero, 2004, for reviews).

Today, some authors declare empathy to be one of the most investigated topics in social neuroscience and other domains (Jacobs, 2012). However, hand in hand with the multitude of approaches from several disciplines the number of proposed definitions and conceptualizations of empathy is constantly growing (Vignemont & Singer, 2006). Thus, before attempting to investigate the topic of empathy it is necessary to discuss the most important of these definitions with regard to their strengths and shortcomings. This discussion is an important framework for choosing an operational definition that optimally suits the present investigation.

2.1.1 Definition

According to Baron-Cohen and Wheelwright (2004), in the last century the concept of empathy was mainly addressed within two main research traditions: an affective and a cognitive approach. Within both approaches the multiple definitions vary in how broad versus narrow they conceptualize empathy.

The affective approach understands empathy as an emotional response to the emotions of another individual. An example for a broad definition within the affective

approach may be Hoffmann's (2000) conceptualization of empathy as any affective response that is more appropriate to the situation of another person than to one's own. This may encompass conscious and unconscious emotions, basic and complex emotions, etc., independently of higher-order cognitive processes like perspective taking or the evaluation of the situation. The advantage of such a broad definition is that it subsumes many relevant phenomena of interest to research on empathic abilities, and important for behavioral outcomes of empathy. However, broad definitions deny precise claims about the nature of empathy because they do not clearly differentiate between empathy and related mechanisms as theory of mind, emotional contagion, compassion, and others. Further, differences in empathic abilities and behavioral outcomes of empathy cannot be explained accurately with regard to concrete affective processes.

For this reason, some authors prefer narrower definitions of empathy. For example, de Vignemont and Singer (2006) propose four distinct criteria that define empathy: (a) that a person is in an affective state, (b) that this state is isomorphic and equivalent to the affective state of another person, (c) that the affective state is the outcome of the observation or imagination of the affective state of another person, and (d) that the person is aware of the fact, that the other person is the source of one's own affective state. This definition enables to differentiate precisely between empathy and similar constructs. For example, theory of mind or cognitive perspective taking does not meet the first criterion (a), because it does not necessarily involve affection. Emotional contagion does not meet criterion (d), because this process can also happen unconsciously, as in infants or animals. Compassion or sympathy does not meet criterion (b), because the affective reaction within compassion may be different than the observed affective state (e.g., one could feel horrified when observing someone in pain). For an elaborated discussion of the differentiation between empathy and similar constructs, see also Section 2.2. De Vignemont and Singer (2006) provide investigations on the neuronal basis of empathy that support their narrow definition of this construct.

However, it is also possible to define empathy in more cognitive than affective terms. In contrast to the affective approach, cognitive theories regard empathy as the understanding of the emotions of another person (Baron-Cohen & Wheelwright, 2004).

Hogan's definition of empathy as the apprehension of the mental state of another person without experiencing that person's feeling (Hogan, 1969) may serve as an example for a broad conceptualization within the cognitive approach. However, broad definitions here are often equivalent with common definitions of theory of mind, perspective taking, or mindreading (e.g., Astington, Harris, & Olson, 1988; Sodian & Kristen, 2010). Thus, the incremental validity of such operationalizations of empathy may be questionable.

A narrower definition within the cognitive approach was used by Wispé (1986) to differentiate between empathy and sympathy. Wispé understands empathy as an attempt of a self-aware being to comprehend positive and negative experiences of another self-aware person without judgements. She defines several important conditions for empathy: It depends upon imaginal and mimetic capabilities, it is most often an effortful process, and it has often the purpose of providing understanding for at least one of the involved parties. Further, she labels sympathy as a form of relating and empathy as a form of knowing. By creating a link between empathy, the intention to understand the other person, and a non-judgemental attitude, Wispé provides a more elaborated framework that separates empathy from general cognitive abilities and points to its specific social functions. However, it may be debatable whether also affective and not only cognitive mechanisms play an important role in comprehending positive and negative experiences of another person.

As stated before, it would be possible to discuss countless further definitions of empathy. Nowak (2011) provides 57 different scientific definitions of the term empathy in her dissertation, demonstrating the divergence and disagreement on this issue. However, the benefit of a further in-depth discussion of all these possible conceptualizations of empathy for the present investigation is limited. Given the background of both presented approaches towards the understanding of empathy, the next appropriate step is finding an operational definition for this study.

Today, many researchers agree upon the fact, that cognitive and affective facets of empathy are not mutually exclusive, but complementary. Thus, the two primary traditional approaches towards understanding empathy have been merged into one very popular conceptualization of empathy: the ability to understand and to share the emotions of others.

This broad definition is accepted and being used by many researchers (e.g., Baron-Cohen & Wheelwright, 2004; Davis, 1994; Decety & Jackson, 2006; Dziobek et al., 2008).

Because the present investigation deals neither exclusively with affective nor exclusively with cognitive empathic abilities within social groups, a definition of empathy encompassing both affective and cognitive facets appears appropriate for this investigation. For this reason, empathy is defined as the ability to understand and to share the emotions of others in the present study. However, to allow a precise differentiation between empathy and related constructs, it is important to add one additional criterion to this definition: In this investigation, empathy is defined as a conscious process, which means that the person who understands and shares the emotions of another person is fully aware of this process and the fact, that the other person is the source of the own experience.

2.1.2 The Multidimensional and Multifaceted Structure of Empathy

A conceptualization of empathy that subsumes related cognitive and affective phenomena can be regarded as a multidimensional approach towards the understanding of empathy. More precisely, this approach comprises two separate, but related dimensions of empathy: cognitive empathy - the ability to understand the emotions of other persons - and emotional empathy - the ability to share the emotions of other persons (Davis, 1983; Davis, 1994; Dziobek et al., 2008).

Subsequent research on these dimensions revealed evidence for different ontogenetic trajectories (Singer, 2006) as well as different phylogenetic trajectories (Preston & de Waal, 2002) underlying cognitive and emotional empathy. Further, both dimensions differ regarding several psychological disorders (Blair, 2005) and may also be related to different neural systems (Decety & Jackson, 2004) and different behavioral outcomes (Smith, 2006). Thus, it is interesting to highlight the most important specifics of both postulated dimensions of empathy.

Cognitive empathy can be regarded as a specific subcomponent of the general mentalizing ability that humans use to explain and predict their own behavior and the behavior of others by attributing mental states to them (Decety & Jackson, 2004). Most

researchers agree that the ability to understand the emotions of others involves advanced cognitive perspective-taking systems (e.g., de Waal, 2008; Shamay-Tsoory, Aharon-Peretz, & Perry, 2009). However, cognitive empathy is also dependent on rather basic cognitive processes and functions as attention and working memory. These processes and functions are important to note and process situational cues in social interaction (Rankin, Kramer, & Miller, 2005). Further, cognitive empathy entails the ability to recognize and interpret facial expressions, voice prosody, social salience of movements and eye region expressions (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). The recognition and processing of such relevant social cues provides relevant information on the emotional state of another person and is the basis of higher-order cognitive processes involved in empathy like abstract reasoning or cognitive flexibility (Rankin et al., 2005). Within these higher-order processes complex meta-cognitive techniques, as for example shifting one's attention back and forth to compare and contrast the experience and behavior of another person to one's own cognitive and emotional state, can be used to provide an adequate empathic response. Figure 1 subsumes the relevant functions and processes that are involved within cognitive empathy. All relevant functions are interdependent and can influence each other. However, higher-order processes are reliant on basic cognitive processes. For example, abstract reasoning on the emotional state of another person is not possible without paying attention to the emotional state of this person or without providing the required working memory capacities.

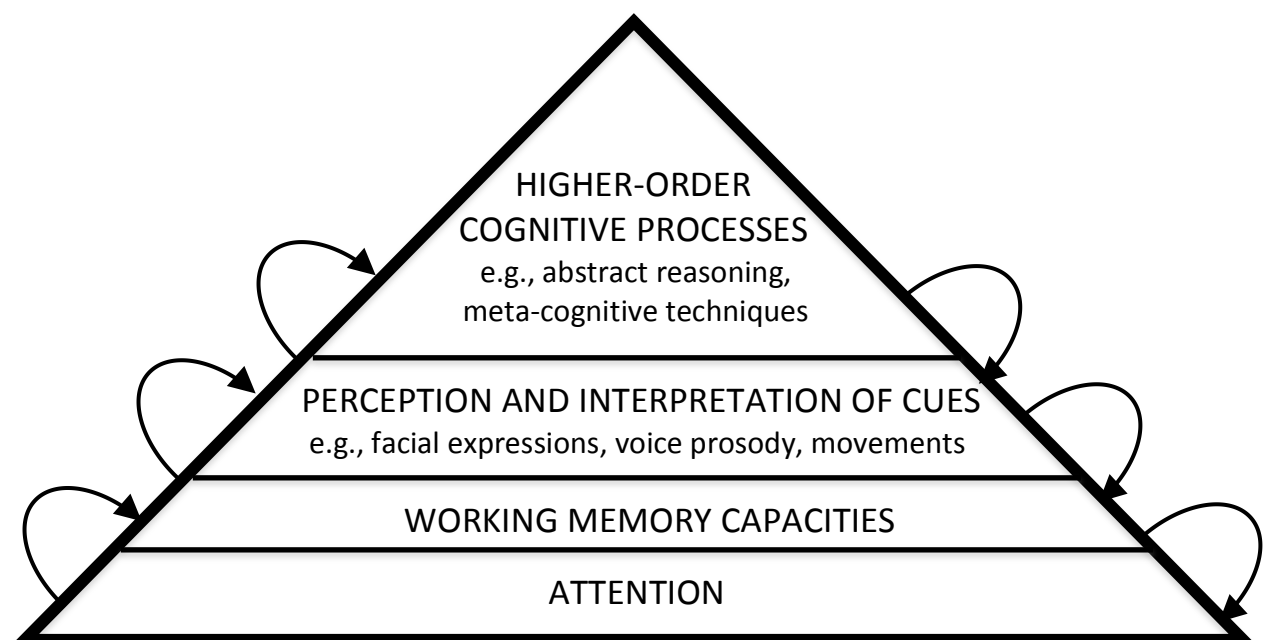


Figure 1. Interdependence of Cognitive Functions and Processes within Cognitive Empathy

Emotional empathy can be defined as the ability to detect the immediate emotional state of another person. According to Preston and de Waals' (2002) perception-action hypothesis, observing the behavior of another person automatically elicits an activation of one's own representations of this behavior. This process stands at the beginning of the empathic experience and seems to be associated with functions of the already mentioned mirror neuron system (Nummenmaa, Hirvonen, Parkkola, & Hietanen, 2008). This elementary emotional response is usually defined as emotional contagion and can be also observed in infants and many animal species (de Waal, 2008; Preston & de Waal, 2002). Further, emotional empathy is dependent on the emotional arousal system and the basic emotional responsiveness of a person (Rankin et al., 2005), as any emotional experience. The higher the individual's emotional responsiveness, the more probable is an emotional reaction to the emotions of others. However, the mentioned basic processes are necessary, but not sufficient for the experience of emotional empathy: According to the proposed definition of empathy, emotional empathy involves the consciousness that the other person's experience and not one's own is of primary importance. This more complex emotional reaction is often described by the term *empathic concern* (Davis, 1983; Rankin et al., 2005). Figure 2 depicts the relevant functions and processes involved in emotional empathy.

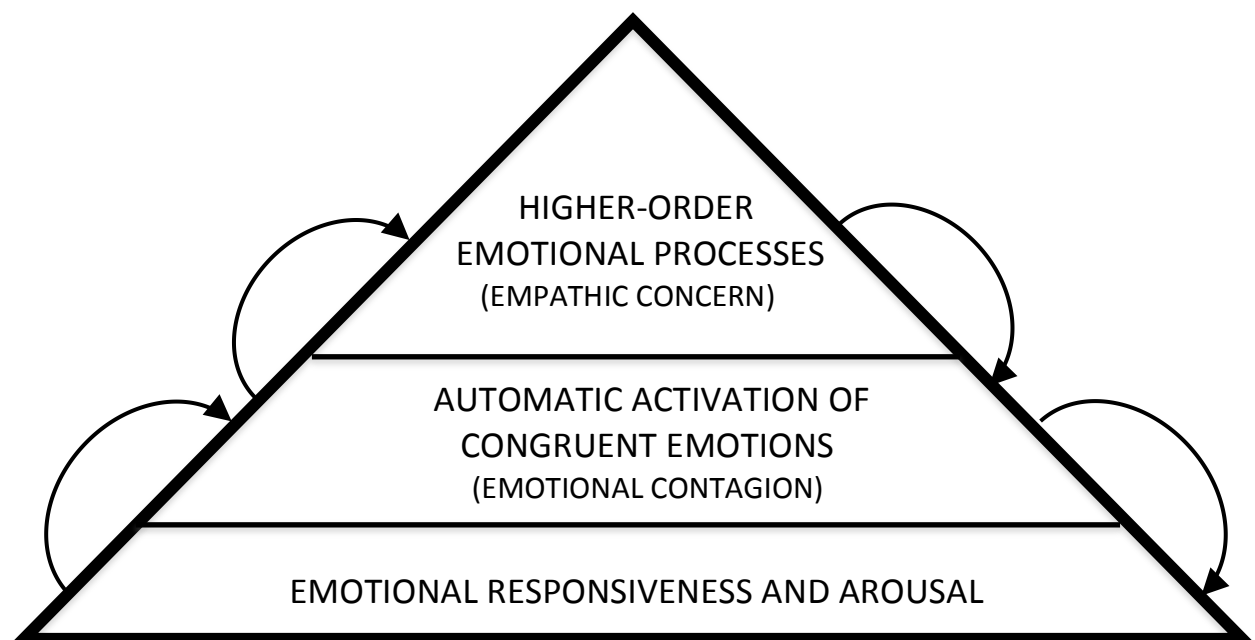


Figure 2. Interdependence of Emotional Functions and Processes within Emotional Empathy

Some authors use the term empathic concern to describe not only the experience of identical emotions but also further other-oriented emotions as tenderness, sorrow, or grief (e.g., Batson, 2011). These conceptualizations are comparable to commonly used definitions of compassion and sympathy and will be discussed in detail in Section 2.2.3.

After having discussed the characteristics of the two proposed dimensions of empathy including their associated facets it is possible to ask a critical question: Is it really necessary to combine both dimensions to understand the nature of empathy – or would it be rather suitable to define them as two separate, independent constructs, e.g., as perspective taking and empathic concern?

Present psychological research demonstrates that despite of their different ontogenetic and phylogenetic trajectories cognitive and emotional empathy are strongly interdependent (e.g., Decety & Jackson, 2004; Lamm, Batson, & Decety, 2007; Preston & de Waal, 2002; Smith, 2006). The most simple and suitable explanation for this interdependence is that understanding an emotional state of another person usually precedes consciously sharing this emotional state. Thus, the ability to share the emotions of others is dependent on the ability to understand the emotions of others (Decety & Jackson, 2004). On the other hand, emotional contagion, which is a basic form of emotional empathy and usually occurs unconsciously and automatically, can also help to understand and to elaborate the emotions of another person cognitively (Shamay-Tsoory et al., 2009).

Thus, both dimensions depend on each other and should be regarded as two separate, but interdependent systems of empathy (Figure 3).

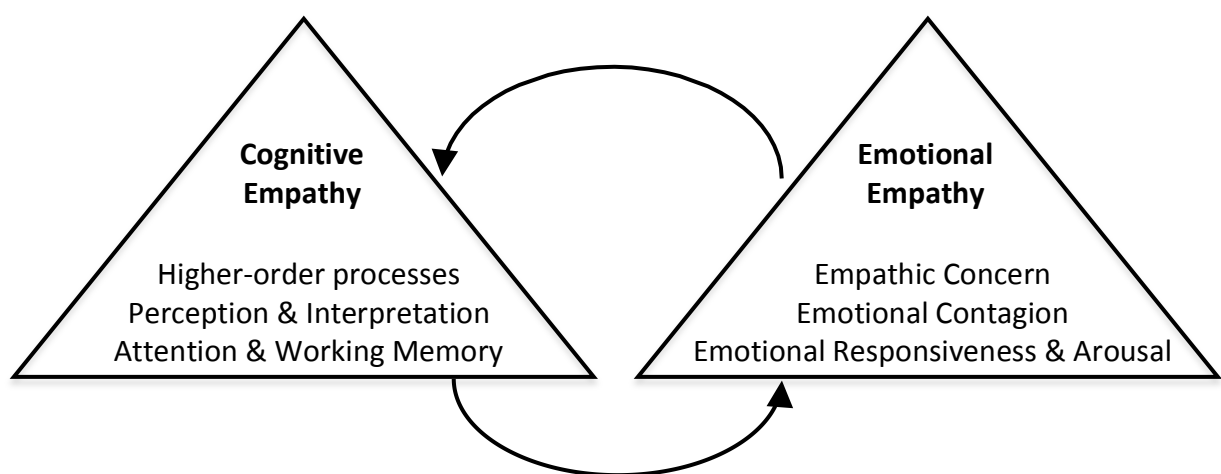


Figure 3. Interdependence of Cognitive and Emotional Empathy

After discussing the facets of the two postulated dimensions of empathy and their interdependence, two additional popular approaches that further differentiate the multidimensional structure of empathy should be introduced.

One of the most popular operationalizations of empathy was conducted by Davis (1980; 1983; 1994). This approach is also based on the discussed dimensions of cognitive and emotional empathy. However, cognitive empathy is further differentiated into two subdimensions: perspective taking and fantasy. Davis (1983) defines perspective taking as the ability to spontaneously adopt the psychological point of view of others and fantasy as the ability to transpose oneself imaginatively into the feelings and actions of others. Alike, emotional empathy is differentiated into two further subdimensions: empathic concern and personal distress. Empathic concern is defined as the disposition to have other-oriented feelings of sympathy for unfortunate others and personal distress as the disposition to have self-oriented feelings of personal anxiety in interpersonal situations (Davis, 1983). This conceptualization of empathy was the basis for the development of the Interpersonal Reactivity Index (IRI; Davis, 1980; 1983) that is also used within the present investigation (see also Section 4.3.1.1). Davis (1983) demonstrated that each of the four subscales is distinctively interrelated with external criteria as social functioning, self-esteem, and emotionality and also with previous empathy measures. However, a limitation of this operationalization may lie in the unclear differentiation between empathy and related constructs. For example, empathic concern as defined by Davis (1980; 1983) is to a great extent concordant with prevalent definitions of compassion (Gilbert, 2005). Further, general perspective taking is often differentiated from empathy and regarded as an independent construct (e.g., Batson, Early, & Salvarani, 1997). Finally, the operationalization of empathy within the IRI (Davis, 1980; 1983) does not consider the already mentioned basic functions that are associated with cognitive empathy (e.g., working memory, attention) and emotional empathy (e.g., general emotional responsiveness; Rankin et al., 2005).

In summary, the IRI (Davis, 1980; 1983) represents an important additional conceptualization of empathy. However, due to certain limitations it should not be used as an exclusive conceptualization.

Another important multidimensional approach towards the understanding of empathy was presented by Batson (2009). This approach distinguishes eight distinct and interrelated phenomena regarding the construct of empathy: (1) knowing another person's internal state, (2) adopting the posture or matching the neural responses of an observed other, (3) coming to feel as another person feels, (4) intuiting or projecting oneself into another's situation, (5) imagining how another is thinking and feeling, (6) imagining how one would think and feel in the other's place, (7) feeling personal distress at witnessing another person's suffering, and (8) feeling for another who is suffering.

This conceptualization of empathy unquestionably subsumes further important facets of the ability to understand and to share the emotions of others. However, in its basic structure it is very similar to the subdimensions that were proposed by Davis (1980; 1983): Concept 4 is highly comparable to the subscale Fantasy, concepts 5 and 6 to the subscale Perspective Taking, concept 7 to the subscale Personal Distress, and concept 8 to the subscale Empathic Concern. Thus, as in the IRI (Davis, 1980; 1983), Batson's (2009) proposed facets of empathy are strongly consistent with related constructs as emotional mimicry (concept 2), emotional contagion (concept 3), perspective taking (concepts 4 and 5), and compassion (concept 8). All these related constructs have commonalities with empathy but demand a more precise differentiation (see Section 2.2).

In summary, two steps are necessary for a profound conceptualization of empathy: First, it is essential to consider all relevant dimensions and facets that are related to the ability to understand and to share the emotions of others. Second, it is important to provide a clear differentiation between empathy and related constructs as theory of mind, emotional contagion, compassion, emotional mimicry, and perspective taking.

However, before proceeding to this second step in Section 2.2, it is important to provide some basic information on the psychological functions of empathy – as these functions are also essential for differentiating empathy from related constructs.

2.1.3 Functions of Empathy

The most important functions of the ability to understand and to share the emotions of others can be divided into two categories: mental functions and behavioral functions.

Redmont (1989) describes several important mental functions of empathy. First, empathy plays a central role in making predictions about others and enhances a general understanding of the mental states of other persons. Thus, it affects decision making in interpersonal contexts and mental attributions of others. Especially ethical decision making and moral judgments are strongly dependent on empathy (Eisenberg & Strayer, 1987; Mencl & May, 2009).

Further, empathy enhances more positive emotions in interactions with others, general subjective well-being, and experiencing the own interactions as more meaningful (Grühn, Rebucal, Diehl, Lumley, & Labouvie-Vief, 2008). Thus, empathy triggers a sense of belonging to others and tightens relational bonds (de Vignemont & Singer, 2006).

As opposed to the two functions explained so far, another important mental function of empathy does not point to other persons, but to ourselves. By "*seeing ourselves in others' reflections of us*" empathy is related to the constitution of our self-concept (Eisenberg & Strayer, 1987, p. 110; see also Section 2.3.4). In other words, by understanding what other persons think and feel about us - the so-called perceived public image - we define ourselves and evolve.

However, most researchers focus on the behavioral functions of empathy. A central behavioral function of the ability to understand and to share the emotions of others is to enhance verbal and nonverbal communication between interaction partners. According to Davis (1994), empathy improves communication both quantitatively and qualitatively. In this regard, cognitive empathy makes the exchange of information more effective and efficient, and emotional empathy enables and motivates the interaction partners to provide more personal information to each other.

Yet, the most important behavioral function of empathy is often seen in triggering and enhancing various forms of prosocial and altruistic behavior (Batson et al., 1991; Batson, O'Quin, Fultz, Vanderplas, & Isen, 1983; Davis, 1994; de Vignemont & Singer, 2006; Preston & de Waal, 2002).

Krebs (1975) conducted one of the first prominent empirical investigations on the interrelationship between empathy and altruistic behavior. Within this study, participants who reacted most empathically towards a performer receiving electrical shocks also tended to behave most altruistically and to help the performer at a cost to themselves. Contrary, participants with lower levels of empathy towards the performer had a weaker tendency to react altruistically.

The interrelationship between empathy and altruism was also addressed in a series of six similar experiments by Batson and colleagues (Batson, Duncan, Ackerman, Buckley, & Birch, 1981; Batson et al., 1988; Batson et al., 1991). In these experiments, a young girl being observed by the participants was receiving electrical shocks. During the experiment, participants had the possibility to switch places with the girl and receive the electrical shocks instead of her. Batson and colleagues (1981; 1988; 1991) explained a higher altruistic motivation to help the girl during the experiment with higher levels of empathy. Based on the results they postulated the prominent "Empathy-Altruism Hypothesis": if a person understands and shares the emotions of another person it is more likely that this person will help the other person, regardless of possible gains or personal losses.

Despite of strong limitations of the studies of Batson et al. (1981; 1988; 1991; e.g., that the level of empathy was not measured or that the effect could also be explained by perceived similarity with the girl, see Cialdini, Brown, Lewis, Luce, & Neuberg, 1997) the "Empathy-Altruism Hypothesis" provided an important framework for subsequent investigations on the interrelationship between empathy, altruistic, and prosocial behavior.

Today, most researchers agree upon the fact that empathy triggers prosocial and altruistic behavioral responses (e.g., Decety & Jackson, 2004; de Vignemont & Singer, 2006; Eisenberg & Fabes, 1990; Preston & de Waal, 2002). Evidence suggests that even related basic functions as emotional mimicry enhance prosocial behavior (Baaren, Holland, Kawakami, & Knippenberg, 2004).

However, a less prevalent and therefore often unlighted psychological function of empathy is the enhancement of antisocial behaviors. Recent research suggests that especially cognitive empathy can also be used to manipulate others and maximize one's

benefits at the cost of others (e.g., Galinsky, Maddux, Gilin, & White, 2008; Maddux, Mullen, & Galinsky, 2008). This possible outcome of empathy is reasonable, as the profound understanding of the emotions of others provides possibilities to gain advantage over them. Further, a high level of emotional empathy may also lead to high distress resulting in relational conflicts and conflict escalation (Eisenberg, 2002; LeBlanc, Gilin, Calnan, & Solarz, 2012). These possible negative behavioral outcomes of empathy should be taken into account, when discussing the multiple functions of empathy.

2.2 Distinguishing Empathy from Related Constructs

This dissertation seeks to investigate empathy on an individual and systemic level in relation to specific group parameters. However, after taking the first step towards this aim by conceptualizing and defining empathy, it is crucial to consider the question of convergent and divergent validity of this construct. This issue states the grounding for an adequate and precise operationalization and measurement of empathy.

Empathy shares characteristics with other concepts, e.g. theory of mind (Sodian & Kristen, 2010), emotional contagion (Hatfield et al., 1993), compassion (Gilbert, 2005), emotional mimicry (Hoffmann, 2002), and perspective taking (Walter, 2012). The following sections provide a discussion of the commonalities and differences between empathy and each of these related constructs. An overview of their definitions and differentiating features with regard to empathy can be derived from Table 1.

2.2.1 Empathy and Theory of Mind

Theory of mind is usually defined as the ability to attribute mental states to the persons we are interacting with and to oneself (e.g., Sodian & Kristen, 2010) and therefore a key component of our everyday lives. By understanding the mental states and intentions of other persons, theory of mind enables us to predict their behavior. Therefore it can be regarded as essential to successful social interaction (Leudar, Costall, & Francis, 2004; Repacholi & Slaughter, 2003). The terms theory of mind and mentalizing are often used synonymously (e.g., Baron-Cohen & Wheelwright, 2004; Walter, 2012).

Table 1

Constructs Related to Empathy: Definitions and Differentiating Features

Concept	Definition	Differentiating Features
Theory of mind	The ability to attribute mental states to the persons we are interacting with and to oneself (e.g., Sodian & Kristen, 2010)	Theory of mind is a very broad ability referring to any kind of mental state. Cognitive empathy is a specific subcomponent of theory of mind that refers to emotional states (=affective theory of mind). Emotional empathy and cognitive theory of mind may elicit each other but are associated with functions of rather independent neural networks.
Emotional contagion	The tendency to automatically mimic and synchronize expressions, vocalizations, postures, and movements with those of other persons and the associated emotional conversion with other persons (e.g., Hatfield et al., 1993)	Emotional contagion is a very basal process, which can occur automatically and unconsciously, without knowing that the other person is the source of one's own feelings. Empathy is a conscious emotional and cognitive response that involves higher-order emotional and cognitive processes (e.g., empathic concern or abstract reasoning).
Compassion	Openness to the suffering of others, which involves a desire to relieve suffering, cognitions related to the understanding of suffering, and behaviors that aim towards relieving suffering (Gilbert, 2005)	Common conceptualizations of compassion are very broad and relate to unspecific emotional or behavioral responses to suffering of others (e.g., being despaired when seeing someone in pain). Empathy is restricted to understanding what another person is feeling and sharing the same emotion (e.g., feeling pain when seeing someone in pain). Further, compassion is restricted to negative emotional states. Empathy relates to negative and to positive emotional states.
Emotional mimicry	The automatic imitation of another's facial, vocal, or postural expressions (Hoffmann, 2002)	Emotional mimicry does not necessarily involve any emotional or cognitive reaction to the other's emotional state.

(table continues)

Table 1 (*continued*)

Concept	Definition	Differentiating Features
Perspective taking	The adoption of the mental perspective of another person – or putting oneself mentally in the shoes of the other (Walter, 2012)	Perspective taking is a very broad concept and refers to cognitive and emotional mental states as well as to sensory perceptions (e.g., within visuospatial perspective taking). Cognitive empathy refers only to emotional states. Further, perspective taking does not necessarily involve any emotional reaction in the observer, which is a differentiating feature with regard to emotional empathy.

Even if some authors conclude that the question whether the concepts of theory of mind and empathy overlap or not is an unsolved issue (Decety & Jackson, 2004), most researchers agree that there is some essential conceptual overlap or even congruence between theory of mind and cognitive empathy (e.g., Baron-Cohen & Wheelwright, 2004; Blair, 2005; Dziobek, 2012; Nummenmaa et al., 2008; Walter, 2012). Considering that cognitive empathy is usually defined as the ability to understand and to recognize emotions of others and theory mind as the ability to understand the mental states and intentions of others, the postulated overlap seems reasonable, because an emotion is also a mental state. Subsequently, theory of mind involves also higher-order cognitive processes as abstract reasoning or meta-cognitive techniques (e.g., Santiesteban et al., 2011).

However, these commonalities also point towards one essential difference between both concepts: while theory of mind is a very broad ability encompassing all mental states, cognitive empathy refers to emotional states only. Thus, cognitive empathy can be regarded as a specific subcomponent of theory of mind. Some authors describe the ability to understand the emotions of others also as an *affective theory of mind* (e.g., Kalbe et al., 2010; Sebastian et al., 2012; Shamay-Tsoory et al., 2007; Walter, 2012). Thus, the terms cognitive empathy and affective theory of mind may be used synonymously, at least with regard to the presented definition of cognitive empathy in Section 2.1.1.

Several recent investigations have demonstrated that cognitive theory of mind and affective theory of mind are two rather independent dimensions that differ with regard to neural activation levels and involved neuroanatomical structures (Kalbe et al., 2010), developmental changes (Sebastian et al., 2012), and the level of impairment in specific psychiatric populations (Poletti, Enrici, & Adenzato, 2012; Shamay-Tsoory et al., 2007). For example, negative symptoms of schizophrenia seem to be associated with impairments in the ability to understand emotional mental states, but not with impairments in the ability to understand cognitive mental states (Shamay-Tsoory et al., 2007). On the other hand, patients with cortical (e.g., Alzheimer's disease) and frontal-subcortical dementia show selective impairments in the ability to attribute cognitive mental states while the ability to attribute emotional mental states remains intact (Poletti et al., 2012).

These and other studies point to the importance of a multidimensional operationalization and measurement of cognitive and affective theory of mind (see also Dziobek, 2012). However, as cognitive theory of mind is not a central object of investigation in this dissertation, this discussion should not be further elaborated at this point.

After discussing the commonalities and differences between cognitive theory of mind and cognitive empathy it might be also important to consider the commonalities and differences between cognitive theory of mind and emotional empathy. On a conceptual level, cognitive theory of mind remains neutral to the question whether an emotional state is elicited in the observer, which is a central criterion for emotional empathy. However, it is possible that cognitive theory of mind elicits emotional empathy and vice versa.

An attempt to investigate the interrelationship between cognitive theory of mind, cognitive empathy (or affective theory of mind), and emotional empathy on a neurophysiological level was provided by Walter (2012). By identifying neural correlates for each of the mentioned concepts he demonstrates that cognitive empathy may be an intermediary instance with regard to cognitive theory of mind and emotional empathy: An elicitation of emotional empathy (e.g., for pain) may result in understanding the emotions of another person. This understanding may subsequently lead to mentalizing the person's other mental states, not only the person's emotions. Walter (2012) describes this path as a "low road" or as a bottom-up process. However, the other possibility is that mentalizing on

non-emotional mental states of another person (cognitive theory of mind) may also lead to the understanding of the emotions of this person and finally to experiencing the same emotions (high-road / top-down). As stated before, these three concepts seem to share certain neural networks but also differ with regard to their main neural correlates. For example, cognitive theory of mind is associated with an activation of the dorsomedial prefrontal cortex, cognitive empathy with an activation of the ventromedial prefrontal cortex, and emotional empathy with an activation of the midcingulate cortex (Walter, 2012). Thus, even if emotional empathy may elicit cognitive theory of mind and vice versa, emotional empathy and cognitive theory of mind seem to share less neural networks and be rather independent from each other than cognitive empathy and cognitive theory of mind.

In summary, theory of mind as the ability to infer the mental states of other persons is a very broad concept. Cognitive empathy is a specific subcomponent of theory of mind that may be also described as an *affective theory of mind*. Emotional empathy may elicit cognitive empathy and vice versa. Further, cognitive empathy may elicit the understanding of other mental states than emotions and vice versa. However, even if emotional empathy and cognitive theory of mind are related by this activation path they seem to share less neural networks and be more independent from each other than cognitive empathy and theory of mind.

2.2.2 Empathy and Emotional Contagion

Emotional Contagion can be defined as an emotional state-matching of a subject with an object (de Waal, 2008) or as “the tendency to automatically mimic and synchronize expressions, vocalizations, postures, and movements with those of another person’s and, consequently, to converge emotionally” (Hatfield et al., 1993, p.96). Emotional contagion is usually regarded as a very elementary mechanism and as a basic building block of human interaction (Hatfield, Rapson, & Le, 2008; Singer, 2006).

The adoption of another’s emotional state can occur automatically and unconsciously when experiencing this emotional state with any sensory system. For example, Prehn-Kristensen et al. (2009) conducted an interesting study in which they claim to have induced anxiety by chemosensory stimuli. Participants in an experimental condition were confronted

with sweat of persons who were waiting for an academic examination. Participants in a control condition were confronted with sweat of persons who participated in an ergometer training. Although the chemosensory stimuli were presented with a low intensity (only 50% of participants noticed the smell of sweat) there were highly significant differences between the experimental and control condition in the activation of brain areas involved in processing social anxiety signals (fusiform gyrus), emotional states of others (insula, precuneus), and brain areas associated with attentional control systems (dorsomedial prefrontal cortex, thalamus). Thus, emotional contagion seems to occur even by *smelling the emotions* of other persons.

Further examples for automatic and unconscious emotional contagion may be a mass panic or a group of infants in which one baby starts to cry and the others subsequently also start to cry. In general, emotional contagion is a very ubiquitous and elementary process that can be observed not only in infants but also among various animal species (e.g., de Waal, 2008; Koski & Sterck, 2010; Preston & de Waal, 2002). Wild, Erb, and Bartels (2001) suppose a “prewired” neural basis for emotional contagion as it occurs rapidly, repeatedly, and automatically. Several researchers underline the relevance of the already mentioned mirror neuron system for emotional contagion (e.g., Hatfield et al., 2009; Preston & de Waal, 2002; de Waal, 2008; Singer, 2006). From an evolutionary perspective, emotional contagion is essential for the regulation of social interactions, coordinated activity and cooperation (de Waal, 2008).

In accordance with their definition, Hatfield et al. (1993, 2009) consider emotional mimicry as an underlying mechanism of emotional contagion. However, it might be beneficial to differentiate emotional or motor mimicry from emotional contagion (see Section 2.2.4).

Psychological research suggests that emotional contagion may be a precursor of empathy (Hoffmann, 2000; Singer & Lamm, 2009) or even a crucial component of empathy (e.g., Hatfield et al., 2009). The automatic and unconscious adoption of the emotional state of another person may lead to the understanding of the other’s emotional state and higher-order emotional and cognitive processes.

However, the described characteristics of emotional contagion point towards the crucial difference between emotional contagion and empathy. Empathy is a conscious process that involves higher-order processes as empathic concern for the other person or abstract reasoning on the other's emotional state (see Section 2.1). Thus, empathy relies upon the ability to distinguish whether one's own emotional state or the emotional state of another person is the source of the own affective experience (Decety & Hodges, 2006; Decety & Jackson, 2004; de Vignemont & Singer, 2006; Singer & Lamm, 2009). Infants develop this capacity around 18-24 months of age (Decety & Jackson, 2004). As contagious crying can be observed even in newborn babies (Singer, 2006) it seems reasonable that emotional contagion can occur without the other-self distinction which is necessary for understanding and sharing the emotional states of other persons.

A further differentiation between emotional contagion and empathy can be conducted with regard to the functions of these two mental processes. Emotional contagion has the basic function to enable a quick and immediate behavioral response to the emotional state of the other person to coordinate social interactions and cooperation (de Waal, 2008). However, as emotional contagion does not distinguish between the own emotions and the emotions of the other, the main focus within the process of emotional contagion remains the self. In contrast, empathy is associated with more other-oriented and more complex mental functions as decision making in interpersonal contexts, moral judgements, and enabling us to make predictions considering the behavior of other persons (e.g., Eisenberg & Strayer, 1987; Mencl & May, 2009; Redmont, 1989; see Section 2.1.3).

Summing up, emotional contagion is an elementary mechanism that seems to occur in infants and several animal species. It can occur automatically and unconsciously, without knowing that the other person is the source of the experienced emotional state. Emotional contagion can be a precursor of empathy. However, empathy is a conscious process and involves higher-order emotional and cognitive processes. Further, the psychological function of emotional contagion is rather restricted to enabling immediate behavioral reactions to the emotions of others while empathy is additionally associated with complex mental functions.

2.2.3 Empathy and Compassion

Compassion can be defined as openness for the suffering of other persons, involving a desire to relieve suffering, cognitions related to the causes of suffering, and behaviors that aim towards relieving suffering (Gilbert, 2005, p. 1). The terms compassion and sympathy are often used synonymously (e.g., Batson, 2009; Walter, 2012). There are also some researchers who use both terms in different contexts, suggesting that these two are constructs with a different meaning (e.g., Gilbert, 2005; Singer & Lamm, 2009). However, as these and other researchers do not provide a clear differentiation between compassion and sympathy, both terms will be used synonymously within this investigation.

Common conceptualizations of compassion are very broad and refer to inherently other-oriented emotions, cognitions and behaviors that are a response to the suffering of other persons (Eisenberg, 2000; Singer & Lamm, 2009). Compassion can involve feelings of sorrow or concern for a distressed or suffering other (de Waal, 2008) but also other-oriented altruistic motivations and cognitive processes as perspective taking (Eisenberg, 2000). Compassion is by definition limited to responses to negative emotional states. An example for compassion would be despair and concern for the injured victims of a witnessed car accident.

Baron-Cohen and Wheelwright (2004) assume that compassion is a clear instance of emotional empathy. Indeed, feeling sorrow or concern for the other person is a higher-order emotional process that may be activated by sharing the emotions of the other person and a key component of emotional empathy (see Section 2.1.2). Additionally, as compassion is associated with cognitive processes as perspective taking, it also entails elements of cognitive empathy (Baron-Cohen & Wheelwright, 2004; Walter, 2012). Another commonality is that both empathy and compassion require a self-other distinction (Walter, 2012).

However, the main feature differentiating between compassion and empathy is the affective isomorphism within empathy (de Vignemont & Singer, 2006; Singer & Lamm, 2009; Walter, 2012). While compassion entails any emotional response to the suffering of another person (Eisenberg, 2000), empathy relies on sharing the emotions of other persons, in other words on having actually the same emotion. With regard to the mentioned example for

compassion, empathy would be not to feel despair and concern but to feel the same pain that the injured victims of the car accident feel. Singer and Lamm (2009) describe the difference between empathy and compassion also by the following comparison: Compassion means feeling for somebody while empathy means feeling with somebody.

Another differentiating feature is that compassion is by definition limited to responses to negative emotional states (e.g., Gilbert, 2005) while most operationalizations of empathy entail the understanding and sharing of both negative and positive emotional states (e.g., Dziobek et al., 2008).

Finally, another difference between compassion and empathy stems from the usually very broad conceptualizations of compassion that also entail other-oriented motivations and behaviors (e.g., Eisenberg, 2000). Specific other-oriented motivations and behaviors may be associated with empathy (e.g., Batson et al., 1991; Preston & de Waal, 2002; see Section 2.1.3) but cannot be part of its conceptualization, as the understanding and sharing of the emotions of other persons can elicit various motivations and behaviors, also antisocial behaviors (Galinsky et al., 2008; Maddux et al., 2008).

In summary, compassion entails higher-order emotional and cognitive processes as empathic concern and perspective taking. These processes are also elements of emotional and cognitive empathy. However, affective isomorphism, the understanding and sharing of positive emotional states and an exclusion of motivations and behaviors from the conceptualization are features of empathy that clearly differentiate between both concepts.

2.2.4 Empathy and Emotional Mimicry

Emotional mimicry can be described as the automatic imitation of another's facial, vocal, or postural expressions (Hoffmann, 2002). Some authors use also the term *motor mimicry* to describe this phenomenon (e.g., de Waal, 2008; Hatfield et al., 2009; Singer & Lamm, 2009). Emotional mimicry and emotional contagion have many similarities: Both can occur automatically and unconsciously and can be observed in small infants as well as in various animal species (de Waal, 2008; de Waal & Preston, 2002). Thus, the main differentiating features between empathy and emotional contagion may be also used to

distinguish between empathy and emotional mimicry (e.g., a self-other distinction, see Section 2.2.2).

Some researchers assume that emotional mimicry is a precursor of emotional contagion and empathy (de Waal, 2008; Hatfield et al., 2009). By mimicking the emotions of other persons it is likely to converge emotionally and to engage in higher-order cognitive and emotional processes on the other person's emotional state. Based on a fMRI study, Leslie, Johnson-Frey and Grafton (2004) propose even a *motor theory of empathy*, suggesting that the generation of emotional facial expressions within emotional mimicry and the recognition of these emotions in other persons within cognitive empathy are both associated with the activation of the same premotor neurons.

However, even if emotional mimicry may be a precursor of empathy, the main differentiating feature between both concepts is that emotional mimicry is by definition limited to observable behavior and does not refer to any cognitive or emotional processes. In contrast to emotional mimicry, empathy relies on both cognitive understanding of the emotions of other persons and the experience of identical emotions.

Proposing that emotional mimicry automatically and always leads to an emotional conversion (Hatfield et al., 1993) may be a too far-reaching assumption. Most authors conclude that emotional mimicry does not necessarily involve emotional matching with the other person (e.g., Singer & Lamm, 2009; Walter, 2012). An example for emotional mimicry without sharing the emotional state of another person is contagious yawning (de Waal, 2008). When we see someone yawning who is tired, most of us automatically start to yawn, too, also without being or getting tired ourselves.

Thus, emotional mimicry occurs when a person automatically adopts the emotional behavior of another person with or without emotional matching. Emotional contagion occurs when a person automatically adopts the emotional state of another person (Walter, 2012). Empathy occurs when the emotional matching occurs consciously and when higher-order emotional and cognitive processes are involved.

2.2.5 Empathy and Perspective Taking

Walter (2012) defines perspective taking as the adoption of the mental perspective of another person, or as putting oneself mentally in the shoes of another person. Even if some authors do not differentiate between perspective taking and theory of mind (Singer, 2006), most researchers define these two concepts differently.

The adoption of the perspective of another person requires more than focusing the attention on the other. It is necessary to take the other's perspective in visual, conceptual, and affective domains into account (Ruby & Decety, 2004). Thus, visuospatial perspective taking is also a central subcomponent of perspective taking (Walter, 2012). Therefore perspective taking can be regarded as a mental process that goes further than theory of mind: Theory of mind can be regarded as the ability to understand the mental states of others (e.g., Sodian & Kristen, 2010). Perspective taking means not only to understand but to *experience* the mental states of others. However, the commonalities and differentiating features between perspective taking and empathy are basically the same as between theory of mind and empathy (see Section 2.2.1).

Walter (2012) regards perspective taking as a cognitive mechanism that is important for both cognitive and emotional empathy. Certainly, the ability to adopt the perspective of another person is crucial for understanding and sharing the emotions of others.

The main difference between perspective taking and empathy is that perspective taking can refer not only to the adoption of emotions but also to the adoption of cognitions or visual perceptions.

Ruby and Decety (2004) demonstrated that perspective taking referring to emotions of others activates different neural networks than perspective taking referring to cognitions of others. A similar study was conducted by Schnell, Bluschke, Konradt, and Walter (2011). In this investigation, visuospatial perspective taking was associated with different neural networks (temporoparietal junction) than the adoption of another's emotional state (dorsomedial prefrontal cortex, anterior superior temporal sulcus, temporal poles). These results are in concordance with the findings on affective and cognitive theory of mind and the functional differentiation of these two subcomponents.

Further, perspective taking can occur without eliciting an emotional state in the observer (Walter, 2012). Thus, empathy usually involves perspective taking, but perspective taking does not always involve empathy (de Waal, 2008).

2.3 Trajectories of Empathy

Psychology as a scientific discipline provides certain scientific paradigms to describe, explain, predict, and influence human experience and behavior. A scientific paradigm can be defined as a bundle of theoretical guidelines, questions, and methods that are supposed to provide answers on these questions (Asendorpf & Neyer, 2013, p.5). It is usually accepted by a large number of scientists of a specific scientific discipline and persists over a longer period of time within the historical development of this discipline.

Empathy can also be described, explained, predicted, and influenced based on certain psychological paradigms. For example, the concept of empathy can be investigated from an evolutionary perspective (de Waal, 2008), an ontogenetic perspective (Grühn, Rebucal, Diehl, Lumley, & Labouvie-Vief, 2008), a neuroscientific perspective (Decety & Jackson, 2004), or a dynamic-interactionistic perspective (Trevvarthen & Aitken, 2001).

For a profound understanding of empathy it is crucial to consider and to combine all these perspectives. These research paradigms are especially important to comprehend the existing gender and age differences in empathy. Moreover, every research paradigm provides specific guidelines for the development of methods to influence empathy and is therefore important for possible practical applications of the results of the present investigation.

Out of these reasons, the following chapter provides a short overview of the most relevant research perspectives on empathy and its trajectories.

2.3.1 Phylogenesis

Unquestionably, empathy has a genetic component (de Waal, 2008). According to the postulates of evolutionary psychology human psychological traits as empathy are evolved adaptations that were favored by natural selection (Asendorpf & Neyer, 2013; de Waal, 2008). Evolutionary psychologists distinguish two types of explanations for the existence of a psychological trait: Ultimate causes, which relate to the fitness consequences of a trait or

why the associated genes were favored by natural selection, and proximate causes, which relate to the concrete mechanism that enables the experience or behavior (Asendorpf & Neyer, 2013; de Waal, 2008; Preston & de Waal, 2002).

The analysis of the ultimate and proximate causes of empathy and cross-species comparisons may provide answers on the question why we are able to understand and to share the emotions of others and important insights with regard to the role of empathy in social groups.

De Waal (2008) assumes that the phenomenon of emotional contagion is phylogenetically continuous and as old as mammals and birds. Several experiments have demonstrated that rats, mice, rhesus monkeys, crows, ravens, and many other animal species react with observable distress to another conspecific in distress or pain (de Waal, 2004; de Waal, 2008; Hatfield et al., 2009; Preston & de Waal, 2002). Further investigations suggest that great apes as chimpanzees, bonobos, gorillas, and orang-utans are even capable of higher-order emotional and cognitive processes that are associated with empathy as cognitive and affective theory of mind (O'Connell, 1995; Povinelli, 1996; Suddendorf & Whiten, 2001) and empathic concern (de Waal, 2008; Hirata, 2009; Mason & Bartal, 2010; Romero, Castellanos, & de Waal, 2010).

These investigations provide at least indirect evidence for the assumption that most animal species are able to converge emotionally with other animals (Plutchik, 1987) and that some higher animal species as great apes seem to be able to distinguish between own emotional states and the emotional states of others, and therefore to understand and consciously share the emotions of others (Thompson, 2001). This evidence suggests that unconscious emotional contagion may not only serve as a precursor mechanism of empathy in general (see Section 2.2.2) but that it may also be an evolutionary precursor of empathy.

De Waal (2004; 2008) addressed this phylogenetic continuity by establishing a so-called "*Russian Doll Model*" of empathy. This model suggests that elementary mechanisms as unconscious emotional state matching underpin more complex and advanced components of empathy as empathic concern and perspective taking. The latter mechanisms depend on prefrontal activity and an increasing self-other distinction, but they remain

connected to the core of any empathic response: a simple perception-action mechanism, which enables the organism to copy the behavior and emotional state of another organism. This “*hard-wired socio-affective basis*” (de Waal, 2008, p. 288) enables animals and humans to coordinate their activities, to cooperate and to pursue shared goals – the more an animal is capable of complex, higher-order cognitive and emotional empathic processes, the more complex and advanced forms of cooperation and goals are possible (de Waal, 2008; see Section 2.4.3).

However, after describing the phylogenetic continuity of empathy, it is important to ask for what reasons we are able to understand and to share the emotions of others considering the evolutionary perspective on this phenomenon. Hence, ultimate and proximate causes of empathy should be taken into consideration.

The already described link between empathy and the coordination of activities and cooperation can be regarded as a central ultimate cause for the evolution of empathy (de Waal, 2008; Preston & de Waal, 2002). From a phylogenetic perspective, the genes of a highly cooperative group of organisms have higher survival chances than the genes of a less cooperative group of organisms.

Plutchik (1987) similarly describes ultimate causes for the evolution of empathy by providing concrete examples: emotional responses to the emotional states of other living beings have survival values because they enable the organisms to gather and hunt for food as well as to respond appropriately and quickly to common threats as predators or natural hazards, e.g. wildfires or storms.

Another central ultimate cause for the phylogenesis of empathy can be used as an explanation for the striking gender differences in empathic abilities that are often found within empirical investigations: Empathy has also the function to bond individuals to one another, especially mothers to their infants (Plutchik, 1987). Simply put, any mother who was not able to recognize and share the emotional states of her infant (e.g., pain or distress because of hunger) was putting her infant’s life at risk. The lower the empathic abilities of the mother, the higher the probability that the child died. If the child died, the genes of the less empathic mother were eliminated from the gene pool.

This evolutionary mechanism provides a simple and excellent explanation for the findings that women outperform men in theory of mind tasks (e.g., Baron-Cohen & Wheelwright, 2004), are more susceptible to emotional contagion than men (Doherty, Orimoto, Singelis, Hatfield, & Hebb, 1995), have higher levels of cognitive and emotional empathy than men (e.g., Lennon & Eisenberg, 1987; Toussaint & Webb, 2005) and even for gender differences in neural correlates of mentalizing and empathic abilities (e.g., Derntl et al., 2010; Krach et al., 2009). Certainly, gender stereotypes and other cultural influences may also account for the mentioned gender differences (see Section 2.3.4), but it is also possible that the described phylogenetic background enhanced the cultural evolution of these social learning prerequisites.

From an evolutionary point of view, the proximate mechanism explaining the phenomenon of empathy is the already mentioned mirror neuron system (e.g., Preston & de Waal, 2002). This neuronal system enables an organism to react to the perception of another organism's state with an activation of corresponding neural representations. This system and the neural basis of empathy will be further discussed in Section 2.3.3.

Further, empathy itself can be regarded as a proximate mechanism for altruism and prosocial behavior (de Waal, 2008; see Section 2.1.3).

In other words, you may help another person in pain because you understand and feel the pain – and you may understand and feel the pain of this person because of the mirror neuron system.

2.3.2 Ontogenesis

The ontogenetic trajectories of empathy are similar to the phylogenetic trajectories that have been described in the previous section.

Newborn babies respond significantly more with crying when they hear another baby cry (Singer, 2006). This effect does not occur when newborns hear the sound of white noise, their own cry, or a synthetic cry (Decety & Jackson, 2004). As newborns are not able to adopt the perspective of another person or to distinguish their own emotional states from the emotional states of others (e.g., Bischof-Köhler, 1991), this observation supports the

assumption that unconscious emotional contagion is the hard-wired core of empathy on which more complex and advanced empathic abilities are built upon (de Waal, 2008).

The mechanism of emotional contagion remains prevalent in the first year after birth (Bischof-Köhler, 1991; Ungerer et al., 1990). However, this automatic emotional response to experienced distress of others seems to become more differentiated in the second half of the first year: When being confronted with a videotape of a distressed peer, some infants tend to ignore this cue, some show self-comforting behaviors as sucking on their fingers or hands, and some become quiet and focus their attention on the peer infant (Ungerer et al., 1990). The authors of this investigation provide evidence for an interrelationship between these individual differences in primitive emotional responding and early self-regulation competencies (measured with the Infant Coping Behavior System by Gianino & Tronick, 1988). Bischof-Köhler (1991) describes this phenomenon as an increased social referencing which occurs in the second half of the first year.

However, most authors agree that more advanced and complex emotional responses than emotional contagion require abilities as self-recognition, self-objectification, person permanence, and finally self-other differentiation. Thus, the next step towards more advanced emotional and cognitive processes is the ability of self-recognition, which usually develops between 18 and 24 months of age (e.g., Decety & Jackson, 2004; Gallup & Platek, 2002). A study investigating the interrelationship between self-recognition and the development of empathic abilities was conducted by Bischof-Köhler (1991). Within this study, the self-recognition of 36 infants (age-range 16 to 24 months) was tested with the rouge-test, and empathic responses were operationalized by observable concern, compassion, and prosocial interventions for a grown-up female playmate with a teddy-bear that lost his arm during the experiment. The results indicated that prosocial interventions as attempts to repair the teddy of the playmate or urging the mother to approach the playmate as well as concerned expressions were highly associated with self-recognition: No child that failed to pass the rouge-test demonstrated one of the mentioned behaviors. Further, even at this young age, girls showed more of these prosocial interventions than boys (compare Section 2.3.1).

Even if Bischof-Köhler (1991) interprets these results as an evidence for the simultaneous onset of empathy and self-recognition, it is questionable if Bischof-Köhler (1991) observed conscious empathy in these infants. First, it is possible that the children were sad themselves because of the broken teddy, e.g., because they wanted to play with it themselves. Second, even here it is possible that the children were affected by emotional contagion and that the attention towards the playmate was not an indicator of conscious empathic concern.

Out of this reason, most authors conclude that conscious cognitive and emotional empathy develops later than self-recognition, because it requires a clear differentiation between the own emotions and the emotions of others and the ability to attribute mental states to others – theory of mind (e.g., Singer, 2006). Even if 18-month old children already demonstrate early reasoning abilities and the ability to distinguish between own preferences and the preferences of others (Repacholi & Gopnik, 1997), full theory of mind abilities develop around the age of four years (e.g., Decety & Jackson, 2004; Eisenberg & Strayer, 1987; Singer, 2006). Whether a child is able to impute mental states to others or not is usually tested with the false-belief task by Wimmer and Perner (1983). Within this task, the following story is told to the children: Maxi has some chocolate, puts it into a blue cupboard and goes out. While he is away, his mother comes and moves the chocolate from the blue into a green cupboard. After describing this scenario the children are asked where Maxi is going to look after the chocolate when he returns. Usually, 3-year-old children give the wrong answer (“in the green cupboard”) while 4-year-olds understand that Maxi will falsely believe the chocolate to be in the blue cupboard and answer correctly (“in the blue cupboard”).

The false-belief task refers to cognitive theory of mind and not to affective theory of mind. However, as these two components of theory of mind are strongly interrelated (see Section 2.2.1) the described findings indicate that emotional and cognitive empathy as defined in the present investigation develops around the age of four years.

In the following developmental stage (age between four and six) the link between empathy and prosocial behavior becomes gradually stronger (Barnett, 1987). Also with regard to this interrelationship a clear gender difference was found: Girls show significantly

more empathy-related prosocial interventions than boys (Barnett, 1987). In later childhood (age between six and ten) the link between empathy and prosocial behavior is being reinforced and moderated by moral principles that develop around this age (Hoffmann, 2000). Further, between six and ten more advanced affective theory of mind abilities enable the child to make more complex inferences about the emotional state of the other than in early childhood (e.g., higher-order beliefs, which means to understand what another person feels with regard to the emotional state of a third person; Singer, 2006).

In youth and early adolescence the ability to understand and to share the emotions of others becomes particularly important – at least with regard to significant peers, as maintaining peer relationships is highly associated with well-being and a crucial factor of identity forming at this age (e.g., Mella, Studer, Gilet, & Labouvie-Vief, 2012). Thus, deficits in empathic abilities in adolescence are strongly interrelated with the occurrence of psychiatric disorders as for example schizophrenia (e.g., Shamay-Tsoory et al., 2007) or antisocial and aggressive behavior at this age (see Lovett & Sheffield, 2007, for a review).

The transition from adolescence to adulthood (around 20) is associated with the onset of elaborated emotion regulation abilities. With regard to empathy, young adults are better capable of regulating their personal distress resulting from sharing the negative emotions of others (e.g., pain), which enables them to provide a more other-oriented and adequate empathic and behavioral response than adolescents (Mella et al., 2012).

In general, psychological research on the development of empathy concentrates on the ontogenetic trajectories in infancy, childhood, and youth rather than on later developmental stages. Thus, research on the development of empathy after adolescence is scarce. Only few studies investigated the ontogenesis of empathy across life span, and these few studies provide inconsistent results. Schieman and van Gundy (2000), for example, reported a decrease of self-reported emotional empathy in late adulthood. On the other hand, Bailey, Henry, and Hippel (2008) identified a decrease of cognitive empathy in late adulthood, but no change in the level of emotional empathy. Finally, Sze, Gyurak, Goodkind, and Levenson (2012) found an increase in self-reported emotional empathy in late adulthood. However, as all of the mentioned studies used cross-sectional designs, Grün et

al. (2008) conducted a longitudinal study spanning 12 years (with participants ranging in age from 10 to 87 years). Within this study, they found no age-related decline in self-reported empathy, which indicates methodological shortcomings (e.g., cohort effects) in the previously mentioned investigations. However, another cross-sectional study by O'Brien, Konrath, Grühn, and Hagen (2013) found quadratic effects of age on empathy across the life span: Middle-aged adults reported higher emotional and cognitive empathy than young and older adults within this study.

In summary, the findings on the development of empathy in adulthood are strikingly inconsistent and demand further investigation.

2.3.3 Neurobiology

The previous sections provided a differentiation between elementary emotional responses as automatic emotional contagion and higher-order empathic processes as empathic concern or affective theory of mind with regard to the phylogenetic and ontogenetic trajectories of these facets of empathy.

Neuroscience provides further evidence for a dissociation of automatic emotional contagion and higher-order empathic responses by identifying distinct brain areas that are associated with both phenomena.

Automatic emotional contagion is associated with the activity of the sensorimotor cortex, which is involved in the processing of sensory experiences and motor functions, and the limbic system (Singer, 2006). The limbic system is a set of brain structures including the hippocampus, the amygdala, gyrus cinguli, and others. It is phylogenetically old and involved in processing of emotional cues, the recognition of facial expressions and body postures, as well as further general functions, e.g. learning and long-term memory (e.g., Dunbar, 1998). Emotional contagion seems to be based on the activity of shared affective neuronal networks within the sensorimotor cortex and limbic system: the already mentioned mirror neuron system (de Vignemont & Singer, 2006; Preston & de Waal, 2002). Even if some recent neuroscientific studies question the existence of mirror neurons in humans (Dinstein, Gardner, Jazayeri, & Heeger, 2008; Lingnau, Gesierich, & Caramazza, 2009), a large number of neuroscientists still regard mirror neurons as the most suitable explanation for automatic

emotional contagion (e.g., Bastiaansen, Thioux, & Keysers, 2009; Nummenmaa et al., 2008; Singer & Lamm, 2009).

In contrast, higher-order processes as affective theory of mind are rather associated with cortical activity of the temporal poles, the posterior superior temporal sulcus, and the medial prefrontal cortex (Decety & Jackson, 2004; Singer, 2006). Specifically, the medial prefrontal cortex is activated when being asked to judge the emotion in another person's gaze, detecting intentions in simple dynamic interaction, or appreciating humor (Decety & Jackson, 2004). The prefrontal cortex is generally associated with complex mental abilities as executive functions, which allow to regulate and to influence psychological processes as attention as well as cognitions and emotions (Decety & Sommerville, 2003). Advanced executive functions related with empathy as the regulation of empathy-related emotional states are probably located in the ventral prefrontal cortex, which is strongly connected to the limbic system (Decety & Jackson, 2004).

This dissociation between the neural substrates of automatic emotional contagion and higher-order empathic processes is consistent with the described phylogenetic and ontogenetic trajectories of empathy, as the limbic system is phylogenetically older than the prefrontal cortex and as it develops before cortical structures in human ontogeny (de Waal, 2008; Preston & de Waal, 2002).

A further elaboration of the neural substrates of empathy may be not necessary, as the present investigation does not use any neuroscientific methods to measure empathy. However, one further neuroscientific study should exemplify that behavioral and self-report measures may not depict the whole truth about empathy. This may be an important limitation with regard to the results of the present study.

Greimel and colleagues (2010) conducted a study on the development of neural correlates of empathy from childhood to early adulthood. Within an fMRI experiment, participants had to infer the emotional state from a presented face (cognitive empathy) and to report their own emotional response to it (emotional empathy). The behavioral performance measures revealed no age-related differences. However, the neural activation of brain areas involved in the processing of social anxiety signals (fusiform gyrus) and

automatic emotional contagion (inferior frontal gyrus) increased with age during the cognitive task, while the neural activation of regions involved in interpreting the intent of others (right precuneus and right intraparietal sulcus) decreased with age during the self-referential task. The authors explain these differences with greater experience during socio-emotional interactions in young adulthood than in childhood, and different cognitive strategies to infer own emotional responses in young adulthood and in childhood (being less dependent on the intentions of others in young adulthood than in childhood).

This speaks for the incremental validity of neuroscientific research on empathy. Unfortunately, an inclusion of neuroscientific methods was not possible in the present investigation out of economical and organizational reasons.

Summing up, neuroscientific research suggests that there is no unitary system for empathy in the brain and that multiple dissociable systems are involved in the experience of empathy (Decety & Jackson, 2004). However, as neuroscience is a young scientific discipline and is still confronted with elementary methodological challenges (Asendorff & Neyer, 2013), the presented findings should be interpreted with caution.

2.3.4 Social Learning

Human experience and behavior develops and is shaped by social interaction with other persons. A prominent model to describe the interdependency between developmental processes and social environments was formulated by Urie Bronfenbrenner (1977; 1979) and consequently elaborated and investigated in later years (e.g., Bronfenbrenner, 1986; Bronfenbrenner & Morris, 1998). Bronfenbrenner (1977; 1979) postulates an ecological embedment of human development: The developing person is influenced not only by immediate social settings as the family (the microsystem) but also by the interrelationships between significant others, e.g. family members or peers (the mesosystem), and larger social contexts that are associated with immediate social interaction partners (the exosystem) and the culture (the macrosystem). However, it is important to consider that the developing person also influences the social environment (see Asendorpf & Neyer, 2013, p.39-48 for various examples).

Thus, individuals seem to develop in a dynamic, continuous, and reciprocal process of interaction with their social environment (Asendorpf & Wilpers, 1998). This dynamic-interactionistic model of human development may be especially important for empathy, as this ability requires social referencing and cannot occur without the presence of others.

Trevarthen and Aitken (2001) underline the importance of intersubjectivity for the development of emotional responses as empathy: As empathy is based on dynamic reactions to another person and dependent on the presence of this person, it can be described as a relational emotional state. It is not only an emotional and cognitive response of a person, but rather an emotional and cognitive response *between* two or more persons (Trevarthen & Aitken, 2001, p. 20). Gieser (2008) describes this relational aspect of empathy within a phenomenological anthropological approach: Empathy describes a mutual, interpersonal process rather than a characteristic of one person and is further based on social interaction, observation, and imitation. Studies on motor mimicry and synchronization, which are associated with empathy (Hatfield et al., 2009; see Section 2.2.4), support the relevance of social contexts for learning to empathize with another person (e.g., Kirschner & Tomasello, 2009). Further, most researchers underline the importance of role-taking during social interactions (especially between the infant and its mother) and reciprocal imitation as an important condition for shared subjective experiences (e.g., Decety & Sommerville, 2003).

Thus, the experience of empathy is shaped by and dependent on the specific relationship and social context in which it occurs.

Thompson (2001) conducted a further interesting philosophical approach considering the role of dynamic interrelation with regard to the development of empathy. According to this model, individual human consciousness is formed by a dynamic interrelation of self and other. Further, the encounter of self and other fundamentally involves empathy. Thus, Thompson (2001) defines empathy as an intersubjective precondition for consciousness. This thesis is supported by the already presented empirical findings with regard to the simultaneous onset of self-recognition and primitive empathic responses in infants (Bischof-Köhler, 1991).

In summary, empathy is an inherently intersubjective process and develops on the basis of social interaction.

As already mentioned in the introduction, psychology as a scientific discipline traditionally provides an individualistic view on human experience and behavior. However, the described dynamic-interactionistic background of empathy demonstrates that this perspective is not sufficient with regard to the ability to understand and to share the emotions of others. The relational and systemic aspects of empathy will be further evaluated in the following chapter.

2.4 Empathy in the Social Context

Section 2.3.4 highlighted the importance of a dynamic-interactionistic view on empathy for understanding this human ability. Empathy cannot be solely regarded as an individual trait of a single person. To understand empathic responses and processes it is necessary to take all persons into account who participate in empathy-related social interactions. Thus, it is important to distinguish between empathy as an individual state or trait and empathy as the state or trait of a relationship or a group. This insight also points to the necessity to develop and apply relational operationalizations within empirical investigations on empathy.

The following section provides an overview over theoretical and methodological differentiations between individual and relational traits as well as possible individual-psychological, relational, and systemic perspectives on empathy. Further, the importance of empathy within social groups will be evaluated, with a strong accentuation of sociological research on group emotions – as well as possible interrelationships between empathy and specific group parameters as diversity and shared goals.

2.4.1 Differentiation Between Individual and Relational Traits

Kurt Lewin, who is often described as a pioneer of research on human relations as well as of psychological research in general (e.g., Dunn, 2011), suggested in his work *“Principles of Topological Psychology”* (Lewin, 1936) that describing and explaining human experience and behavior may not rely on individual traits only. According to him,

psychological research has to additionally consider the social environment and situational influences on human experience and behavior. Within his own investigations, Lewin clearly differentiates between individual-centered and relationship- or group-centered data (e.g., Lewin, Lippitt, & White, 1939). Interestingly, Lewin himself described individual-centered data as *psychological data* and group-centered data as *sociological data* (Lewin et al., 1939).

It seems that Lewin's categorization served as a kind of categorical imperative for both psychology and sociology for decades. Even today, the vast majority of psychological theories and methods remain individual-centered while most sociological theories and methods focus on traits of relationships and groups only (e.g., Asendorpf & Banse, 2000). These disciplinary boundaries exist despite the fact that Lewin demonstrated the importance of considering relationship- or group-centered data within psychology (Lewin, 1936; Lewin et al., 1939), while his colleague Lippitt demonstrated the importance of considering individual-centered data within sociology (Lippitt, 1939).

However, before further discussing the importance and possible applications of relationship- and group-centered methods within psychology, it is important to define and to describe what a social relationship or social group actually is.

A social relationship between two persons can be characterized by stable interaction patterns on the behavioral level and relationship-specific mental representations of the self, the other, and of an interaction script on the cognitive level (Asendorpf & Banse, 2000; Asendorpf & Wilpers, 1998; Baldwin, 1992).

The stable interaction patterns can be described as follows: Within a social relationship there is a probability being significantly higher than random chance that one person will react specifically to a specific behavior of the other person. For example, if two persons have a hostile relationship, it is more probable that they will react in a hostile way to each other. The specific interaction patterns can be partially explained by individual traits, but not only by them. Two persons who are engaged in a hostile relationship have a higher probability to react in a hostile way to each other, even if their individual disposition to show hostile behavior is usually low, as compared to other persons. In other words, there are also interaction variables within social relationships that cannot be explained by individual psychological traits only. These interaction variables within social relationships can be

described as *emergent traits* (Asendorpf & Banse, 2000). Emergent traits arise from the specific interaction history of a relationship, the already mentioned relationship-specific cognitive representations of the involved persons, interaction effects between personality variables, and passive and active partner choice effects (Asendorpf & Banse, 2000).

A social group can be defined as a number of three or more persons who are engaged in dynamic social interaction over a longer period of time (Asendorpf & Banse, 2000). A social group can be also characterized by a specific physical, social, or temporal setting in which the social interaction occurs (Levine & Moreland, 1990). Not all members of a social group are necessarily interrelated to each other within dyadic social relationships. However, the group members are connected to each other by a shared group identity, a “we-feeling” (Myers & Diener, 1995; Seger, Smith, & Mackie, 2009). If the members of social group have a specific common goal, this group can be defined as a *team* (Mathieu, Maynard, Rapp, & Gilson, 2008; McGrath, Arrow, & Berdahl, 2000).

Analyzing emergent traits within social groups is far more complex than within dyads, because there are more levels of possible social interrelationships within groups: The experience and behavior within a social group encompasses dyadic effects between all group-members, but also the effects of triadic interrelationships, etc. (Asendorpf & Banse, 2000). Interaction effects between multiple types of relationships, e.g. a dyadic and a triadic relationship are possible and further complicate the measurement of emergent traits within social groups.

However, some psychologists and sociologists provided impressing approaches towards the measurement of emergent traits both within social relationships and within social groups. Two of these methods should be exemplarily presented at this point to demonstrate the general difference between individual-centered and relationship-centered data.

Gonzales and Griffin (1996; 1997) presented a statistical method to measure emergent traits within a relationship by measuring the latent dyadic correlation between the traits of two persons (Figure 4).

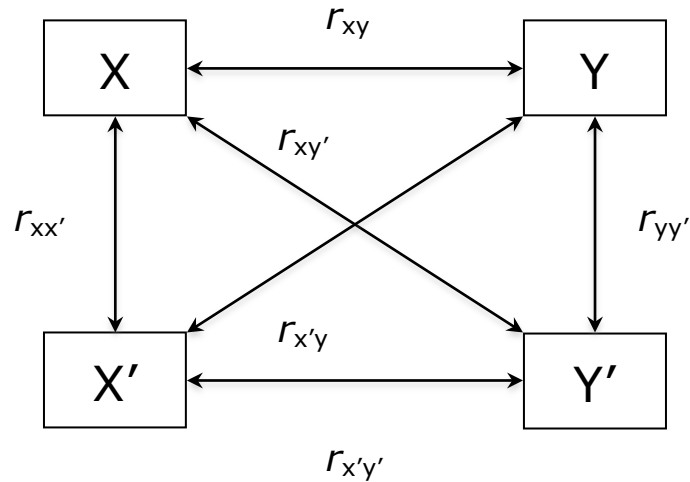


Figure 4. Within-Partner and Between-Partner Correlation Matrix, Based on Gonzales & Griffin (1996; 1997). X = Trait A of Person A; X' = Trait A of Person B; Y = Trait B of Person A; Y' = Trait B of Person B

In this model, the researcher analyses the interrelationship between two variables, e.g. neuroticism and satisfaction with life, within a dyad. r_{xy} and $r_{x'y}$ represent overall within-partner correlations, in this case a general correlation between neuroticism and satisfaction with life on an individual level. $r_{xx'}$ and $r_{yy'}$ represent the pairwise intraclass correlation, in this case the correlation between the neuroticism of an individual with the neuroticism of his or her relationship partner. This measure can serve as an indicator of intra-dyadic similarity (Gonzales & Griffin, 1997). And finally, $r_{xy'}$ and $r_{x'y}$ represent the cross-intraclass correlation, in this example a dyadic effect of one individual's neuroticism on his or her partner's satisfaction with life.

However, cross-intraclass correlations ($r_{xy'}$ and $r_{x'y}$) do not yet represent emergent traits, because they may be influenced by the similarity between the two relationship partners ($r_{xx'}$ and $r_{yy'}$). For this reason, cross-intraclass correlations have to be adjusted for the proportions of variance in X and Y corresponding to shared dyadic effects:

$$r_d = \frac{r_{xy'}}{\sqrt{r_{xx'}} \sqrt{r_{yy'}}}$$

This *latent dyadic correlation* can be interpreted as a direct measure of dyad-level relations between two variables (Gonzales & Griffin, 1997), or in other words, as an emergent trait.

In addition to the presented method for computing dyadic effects, Gonzales and Griffin developed a latent group model (Gonzales & Griffin, 2002), which provides the possibility to assess emergent traits not only on a dyadic but also on a group level by a decomposition and comparison of individual-level and group-level variance.

The latent group model follows the same logic as the latent dyadic correlation: cross-intraclass correlations are adjusted for the proportions of variance in two variables corresponding to shared group-level effects:

$$\sigma_d = \frac{\sigma_{vxy}}{\sqrt{\sigma_{vx}^2 \sigma_{vy}^2}}$$

By this means, it is not only possible to model the personality of a group but also to estimate group-level stabilities over time by computing across-time correlations between group-level latent variables. However, going into details of this complex model would go beyond the scope of this work.

A second approach to measure dyadic effects is the Actor-Partner Interdependence Model (APIM; Cook & Kenny, 2005). This model provides the possibility to measure the development of dyadic effects over time and is based on a conceptual view of interdependence between the interaction partners of a dyad. The crucial feature of the APIM is that the dyad rather than the individual is regarded as the central unit of analysis: The sample size for the analysis is based on the number of dyads and not on the number of individuals. Further, the APIM additionally considers the interaction history of a dyad and embeds the interdependence of two interaction partners into a structural equation model. Figure 5 demonstrates the basic assumptions of this model. Single-headed arrows stand for causal or predictive paths whereas double-headed arrows stand for correlated variables.

As stated before, the emergence of dyadic traits is based upon the specific interaction history of a social relationship. This principle is being focused within the APIM: It relies on the assumption that a person's past behaviors or experiences affect its current behaviors or experiences, but also these of his or her interaction partner.

The paths $X'-Y$ and $X-Y'$ are similar to the described cross-intraclass correlations that were introduced by Gonzales and Griffin (1996; 1997).

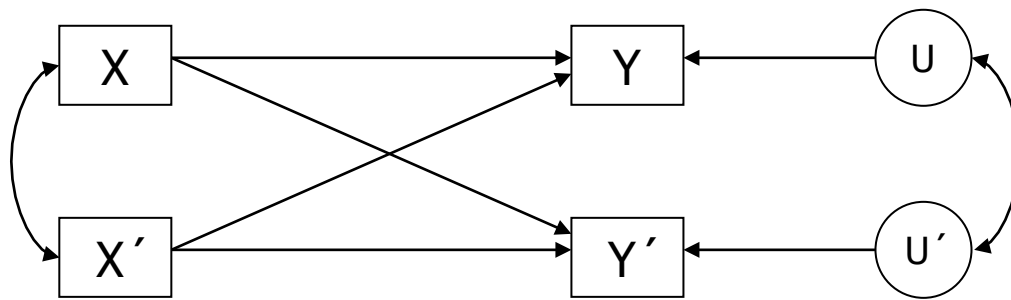


Figure 5. The Actor-Partner Interdependence Model (APIM). X = Data for Person A, Time 1; X' = Data for Person B, Time 1; Y = Data for Person A, Time 2; Y' = Data for Person B, Time 2; U and U' = Residual (Unexplained) Portion of Person A's and B's Time 2 Score

These two exemplary models demonstrate the general difference between an individual-centered and a relationship- or group-centered perspective on human experience and behavior. With regard to the research questions of the present investigation it is important to evaluate the specific implications of these perspectives with regard to the concept of empathy.

2.4.2 Empathy as a Dyadic and Systemic Trait

Although there are statistical methods to investigate the traits of relationships and groups, the application of these methods within psychology is rare. However, the application of dyad- and group-level analyses seems especially important for the investigation of empathy, because of its already described dependency on social interaction and its social functions.

Ickes, Stinson, Bissonnette, and Garcia (1990) conducted one of the first investigations pointing to the importance of differentiating between individual- and dyad-level analyses on empathy. They demonstrated that empathic accuracy, a construct defined as the ability to infer the emotions and thoughts of other persons (nearly congruent with current definitions of theory of mind, see Section 2.2.1), is mainly an emergent product of social interaction processes occurring at the level of the dyad and not at the individual level. Subsequent investigations (Stinson & Ickes, 1992) revealed that dyad-level empathic accuracy is mainly dependent on the acquaintance between both interaction partners.

This work inspired further studies on empathic accuracy as a dyadic trait. Simpson, Ickes, and Blackstone (1995), for example, demonstrated that persons in a romantic relationship tend to have lower levels of empathic accuracy with regard to the emotions of

their relationship partner, when their partner evaluated pictures of unknown and highly attractive opposite-sex persons, with whom the partner had to interact later on. The effect of reduced empathic accuracy was especially strong within close relationships as well as in relationships, in which the partners were insecure about the relationship. Simpson, Ickes, and Blackstone (1995) explain the effect by a *motivated empathic inaccuracy* with the function to protect the self-esteem of a person who is exposed to a situation threatening his or her romantic relationship. Another function of this mechanism may lie in protecting the stability of a romantic relationship.

Within the described investigation, nearly all statistical analyses were conducted on a dyadic level. Unfortunately, the authors did not analyse the difference between dyad-level variables and individual scores. However, another study on empathic accuracy by Ickes et al. (2000a) focused this additional perspective. In this work, Ickes et al. (2000a) tried to determine predictors of empathic accuracy on an individual level. For this reason, they differentiated between variance in empathic accuracy on an individual and a dyadic level. Empathic accuracy on an individual level seems to be dependent on verbal intelligence and interpersonal trust (Ickes et al., 2000a).

A further study by Zaki, Bolger, and Ochsner (2008) differentiated both emotional and cognitive empathy on an individual and on a dyadic level. The authors found that emotional empathy is a significant predictor of cognitive empathy, but only at the level of the dyad, not of the person: The level of emotional empathy of the perceiver predicted cognitive empathy only for expressive targets (for the interrelationship between cognitive and emotional empathy, see Section 2.1.2). This result further demonstrates that empathy as an individual and a dyadic trait are different constructs and that they may rely on diametrically different mechanisms.

However, to the knowledge of the author, there are no empirical investigations operationalizing empathy as a systemic trait, even if the described discrepancies between individual- and dyad-level empathy point to a highly probable discrepancy between individual and systemic empathy. This academic void is remarkable, as especially sociological literature provides a massive amount of positions on the nature and relevance of systemic emotional states and traits.

Barsade and Gibson (1998), for example, describe groups as emotional entities, in which emotions can emerge at the group level and may, but do not have to be experienced by individual members (“top-down” approach). A study by Smith, Seger, and Mackie (2007) suggests that such group-level emotions are distinct from individual-level emotions, depend on the level of group identification, and are socially shared by mechanisms of emotional convergence within groups, e.g. by group emotional contagion. This process of emotional convergence has positive effects on cooperation and task performance and regulates conflicts within groups (Barsade, 2002; Kelly & Spoor, 2006). Further, group-level emotions have the function to regulate intra- and intergroup attitudes and behavior (Kelly & Barsade, 2001; Mackie, Devos, & Smith, 2000; Smith et al., 2007). Some authors even postulate the concept of *intergroup emotions*, which can be understood as a dyadic trait between two systems (e.g., Mackie, Devos, & Smith, 2000; Maitner, Mackie, & Smith, 2006; Paolini, Hewstone, Voci, Harwood, & Cairns, 2006). Intergroup emotions may therefore differ from individual emotions of the group members as well as from dyadic and systemic emotions within a group. However, focusing on empathy not only as a systemic trait but also as an intergroup emotion, a dyadic trait between two systems, would go beyond the scope of this work.

Systemic, group-level emotions can be activated by social categorization cues (Seger, Smith, & Mackie, 2009) and are highly dependent on social norms prescribing emotions and their expression (Barsade & Gibson, 1998), and especially dependent on shared goals (e.g., Kelly & Spoor, 2006).

A simple example from everyday life may illustrate this interrelationship between group emotions and shared goals. Imagine a group of actors planning an outdoor performance in a park. Few minutes before the performance starts they sit together on the grass and have a short break. Suddenly, it starts to rain heavily. What is the difference between the group of actors and other persons who are maybe also relaxing in the park? The emotions that the actors and other persons in the park experience and which are caused by the rain may be the same or at least similar: frustration, disappointment, or even anger. However, the actors experience their emotions as an emotional entity: the rain inhibits their shared goal (the outdoor performance), therefore they are united in their emotional state and probably every actor understands and shares the emotions of the other group members,

as well as the cause of these emotions. In other words, the emotions of the actors emerge on a systemic level, while the emotions of other persons in the park emerge on an individual level.

Even if this example remains on a theoretical level, it illustrates not only the interrelationship between group emotions and shared goals, but also the interrelationship between group emotions, shared goals, and empathy: Shared goals enhance systemic emotional states and traits, and systemic emotions may enhance *systemic* cognitive and emotional empathy within a group.

But what exactly is the difference between individual and systemic empathy? How can both constructs be measured? And what characteristics of a group are empirically interrelated with the level of systemic empathy of this group? These questions have not yet been addressed within psychological or sociological research and represent therefore the main research objectives of the present investigation.

However, before proceeding to the research questions and hypotheses in detail, a brief discussion of the importance of empathy within groups may be helpful to provide some guidelines for the concretion and specification of the research objectives.

2.4.3 The Role of Empathy in Social Groups

As already described in the previous section, emotional convergence within groups enhances cooperation and team performance and decreases conflict potential within these groups (Barsade 2002; Kelly & Spoor, 2006). The transmission of thoughts and emotions seems to be an important condition for any coordinated action of a group (Rafaat, Chater, & Frith, 2009). This function of empathic processes within groups is reflected in several studies documenting an in-group advantage with regard to cognitive and emotional empathy (e.g., Gutsell & Inzlicht, 2010; Thibault, Bourgeois, & Hess, 2006).

These findings are in line with the evolutionary view on the ultimate causes for the emergence of emotional and cognitive empathy in groups that has been addressed in Section 2.3.1: Empathy enhances coordination and cooperation within a group, and a highly cooperative group has higher survival chances than a less cooperative group (de Waal, 2008; Plutchik, 1987). An important proximate mechanism that elicits emotional convergence in

groups may be the immanent need of humans to share their emotions with significant others (see Rimé, 2007, and Rimé, 2009, for a review). A proximate mechanism that promotes cognitive empathy within groups and the successful recognition of emotions between group members may be a high personal relevance and effort in decoding the emotions of members of a social group that individuals identify with. Thibault et al. (2006) demonstrated this effect by creating in- and out-groups within an experiment, in which all participants in all groups shared the same cultural and linguistic background (see the described interrelationship between empathic accuracy and verbal intelligence; Ickes et al., 2000a) and had the same expression styles. Within this experiment, emotional expressions of in-group members were decoded significantly more successful than emotional expressions of out-group members. This result rather cannot be explained by similarity between the members of the in-group, because all participants were similar to each other in central demographical characteristics. Thus, the authors explain the in-group advantage in cognitive empathy by a higher personal relevance and effort in decoding emotional expressions of in-group members (Thibault et al., 2006).

Another effect of empathy in social groups is the enhancement of helping behavior. Stürmer, Snyder, Kropp, and Siem (2006) documented that the often postulated interrelationship between empathy and helping behavior (e.g., Batson et al., 1991; Davis, 1994; Preston & de Waal, 2002) is moderated by group membership: The effect of empathy on helping intentions is stronger when the helper and the target belong to the same cultural group than when they belong to different cultures.

A further central function of empathy in social groups is to regulate intergroup attitudes, emotions, and behavior. Despite the described in-group advantage with regard to emotional and cognitive empathy, empathy with regard to members of other groups seems to be an important condition for cooperation between groups. Empathy seems to be interrelated with harmonic intergroup behavior and the reduction of prejudices, stereotypes, and discrimination (Paolini et al., 2006). Interestingly, even seeing a member of an out-group who displays empathy in front of the perceiver leads to more self-reported liking and less prejudices and stereotypes towards this out-group member than towards out-group members who do not display empathy (Yabar & Hess, 2007; Nadler & Liviatan, 2006).

However, after describing the main functions of empathy in groups, it is important to ask what specific characteristics of a group may be interrelated with systemic empathy. Are there certain group parameters, which enhance or reduce the level of systemic or even individual empathy within a group? The following section will provide an overview over group characteristics that may have a strong impact on systemic empathy.

2.4.4 Interrelationship Between Group Parameters and Empathy

There are several characteristics of a group that are probably interrelated with empathy. Based on theoretical considerations and preliminary findings, possible group parameters interrelated with empathy could be for example group cohesion (Lawler, Thye, & Yoon, 2000; Moody & White, 2003; Rapisarda, 2002; Roark & Sharah, 1989), intermember attraction (Stokes, 1983), social integration within a group (O'Reilly, Caldwell, & Barnett, 1989), the amount and relation of positive vs. negative socioemotional statements within group communication (Keyton & Beck, 2009), amongst many others.

However, all mentioned examples could be both determinants as well as outcomes of systemic empathy. Therefore, it might be more interesting to focus on group parameters that characterize a group from its very beginning and cannot be an outcome of empathic processes within a group.

The theoretical and empirical literature on empathy that was presented in the previous sections mainly suggests two central group parameters that may play a crucial role in regulating and shaping systemic empathy within a group: The (perceived) diversity of a group and initially constituted shared goals. Both parameters and their possible impact on empathy within groups will be discussed in the following subsections.

2.4.4.1 Diversity

Section 2.1.3 highlighted the enhancement of helping behavior as a central function of empathy ("Empathy-Altruism Hypothesis"; Batson et al., 1981; Batson et al., 1988; Batson et al., 1991). However, Cialdini et al. (1997) modified the Empathy-Altruism Hypothesis by identifying *oneness* – the perceived self-other overlap – as a central variable mediating the interrelationship between empathy and helping behavior. In other words, the more similar two persons are, the more probably they will react empathically and with helping behavior to each other.

Even if an ingroup-advantage with regard to emotional and cognitive empathy may occur independently of the similarity between in-group members in demographical characteristics (compare Thibault et al., 2006; Section 2.4.3) the perceived similarity and a related feeling of oneness may have an effect on cognitive and emotional empathy. A finding by Stürmer et al. (2006) supports this assumption: The relationship between empathy and helping behavior towards in-group members is positively associated with perceived similarity to these in-group members.

However, the investigated concepts of oneness and perceived similarity are not precise enough for an operationalization as parameters of a group. Further, Thibault et al. (2006) and Stürmer et al. (2006) investigated both perceived similarity and empathy as individual variables and not as systemic traits of a group. Thus, a concretion and precise operationalization of a systemic group parameter corresponding to oneness and perceived similarity would be helpful.

The concept of group diversity was extensively investigated in social psychology and sociology and could serve as a precisely operationalizable group parameter interrelated with systemic empathy. Diversity can be differentiated in surface-level diversity, which can be defined as heterogeneity among group members in overt, biological characteristics as age, gender, and ethnicity, and deep-level diversity, which can be defined as heterogeneity among group members in attitudes, beliefs, and values (Harrison, Price, & Bell, 1998). Surface-level diversity seems to provoke an initial categorization of other group members that is based on stereotypes, and this initial categorization is in turn associated with perceived similarity or dissimilarity (Harrison et al., 1998).

Even if the interrelationship between diversity and empathy was not yet addressed directly, psychological and sociological literature provides several positions that suggest a dependency between these both variables.

Harrison et al. (1998) demonstrated that surface-level diversity is negatively associated with a group's cohesiveness, even if this effect can be partially neutralized over longer periods of time encompassing intense interactions between group members. Correspondingly, surface-level diversity seems to be associated with lower levels of social

integration (O'Reilly et al., 1989) and higher levels of interpersonal conflict within social groups (Levine & Moreland, 1990). Further, there are multiple findings indicating that teams with high surface-level diversity perform not as good as teams with low surface-level diversity in various tasks, probably because surface-level diversity elicits intergroup biases within a group (e.g., between men and women or between old and young group members) which may disrupt the elaboration of task-relevant information (e.g., van Knippenberg, de Dreu, & Homan, 2004).

The effects of deep-level diversity on a group are far more complex and seem to occur in later developmental stages of a social group (Harrison et al., 1998).

Together with theoretical positions on the interrelationship between the composition of a group and its group-level emotions (e.g., Barsade & Gibson, 1998; Kelly & Barsade, 2001), the mentioned theories and findings suggest a strong interrelationship between the diversity of a group (especially surface-level diversity) and empathy. However, this interrelationship has not yet been addressed within empirical investigations.

2.4.4.2 Shared Goals

A goal can be defined as an internal representation of a desired state (Austin & Vancouver, 1996). This state may be an outcome, an event, or a process. Individual and shared goals within a group or organization may differ (Haas, Sypher, & Sypher, 1992). However, mechanisms of goal contagion enhance the adoption of another individual's goals, if this individual is a member of one's own social group (Loersch, Aarts, Payne, & Jefferis, 2008). Goal contagion seems to be driven by mechanisms of social learning and action observation, which may lead to automatic joint attention, task sharing, and action coordination within social groups (Sebanz, Bekkering, & Knoblich, 2006). Group goals seem to be associated with a significantly higher motivation and performance than individual goals, independently of the ability level (Matsui, Kakuyama, & Onglatco, 1987; Mitchell & Silver, 1990). Individual goals seem to enhance mechanisms of competition, whereas group goals enhance mechanisms of cooperation (van Mierlo & Kleingeld, 2010).

The perception of a shared goal within a social group is associated with higher levels of group identification (Wegge & Hasslam, 2005), group cohesion (Wegge, 2000; Wheelless, Wheelless, & Dickson-Markman, 1982), supportive behaviors towards group members (Aubé

& Rousseau, 2005), willingness to share knowledge with other group members (Chow & Chan, 2008), communication satisfaction (Haas et al., 1992), group performance in various tasks (Kerr & Tindale, 2004; Locke & Latham, 2002; O'Leary-Kelly, Martocchio, & Frink, 1994; Widmeyer & Ducharme, 1997), social trust, and positive affective reactions towards other group members (Klein et al., 2009).

Considering the described evolutionary perspectives on the functions of empathy within groups, specifically regarding the interrelationship between empathy and cooperation and coordinated action within a group (see Section 2.4.3), the presented findings indicate a highly probable interrelationship between shared goals and affective as well as empathic processes within a group.

But what may be the specific principles under which shared goals influence systemic empathy? The example with the troupe of actors in a park that was described in Section 2.4.2 illustrated the possible interrelationship between shared goals and systemic empathy on a theoretical level. An empirical investigation with precise individual and systemic operationalizations of empathy on the association between empathy and shared goals within a group was not yet conducted.

3. Central Research Questions and Hypotheses

The presented theoretical background points towards two central conclusions. First, it is highly important to understand, to operationalize, and to measure empathy not only as an individual trait, but also as a trait of a relationship or of a social system. Second, systemic empathy is probably associated with specific group parameters, especially with surface-level diversity and shared goals of a social group. However, both conclusions have not yet received any attention within empirical academic research, even if such investigations could be associated with enormous implications for applied psychology, for example with regard to the formation and structure of therapy groups, work teams, learning communities, etc. Therefore, the two mentioned conclusions motivate the main research questions guiding the present study:

1. Can empathy be measured empirically as a systemic trait, and what exactly are the differences to empathy as an individual trait?
2. What is the interrelationship between diversity and systemic empathy in social groups?
3. What is the interrelationship between the existence of a shared goal and systemic empathy in social groups?

Providing answers on the first research question requires the development of new, systemic operationalizations of empathy and correspondingly the development of new methods for the measurement of empathy as a systemic trait. To ensure the convergent and divergent validity of these measures, the present investigation is based on an approach that is comparable to the multitrait-multimethod approach (Campbell & Fiske, 1959): Both individual and systemic empathy are each measured with several different methods, to investigate the difference (a) between individual and systemic measures, as well as (b) between different individual measures, as well as (c) between different systemic measures. These different methods entail self-report measures, image-based performance tests, a naturalistic social game, and communication parameters derived from video supported observation. Each of these methods has advantages and disadvantages. For example, self-report measures may be biased by various kinds of cognitive distortions, response-biases, or

social desirability (e.g., Howard & Dailey, 1979; Paulhus, 1984) but offer the most direct assessment of a person's actual thoughts and emotions. Image-based measures provide a higher ecological validity than questionnaires (Dziobek, 2012) but may not fully correspond with actual behavior. On the other hand, communication parameters and social games associated with empathy may not fully correspond with the actual thoughts and feelings of the tested persons but may provide the strongest and ecologically valid implications with regard to the actual behavior of these persons. The only way to combine and to maximize the mentioned benefits and to diminish the disadvantages of these methods is to apply all of them and to investigate the differences between the empirical data obtained by them.

To answer the second research question, small groups of 7 persons were chosen as the central unit of investigation and analysis. This decision has practical reasons: The bigger a social group, the more complex it is to study this group, first because of the multiple levels of social relationships (dyadic, triadic, etc.; see Section 2.4.1) and second because of the possible interactions between these levels.

Further, the construct of diversity was limited to surface-level diversity in demographical factors, deep-level diversity was not investigated. As already stated in Section 2.4.4.1, the level of surface-level diversity is most likely associated with perceived similarity or dissimilarity between members of a social group because it elicits an initial categorization of other persons (van Knippenberg et al., 2004). On the other hand, the effects of deep-level diversity within groups seem to emerge only in later developmental states of a social group and are far more complex than the effects of surface-level diversity (Harrison et al., 1998). Both factors complicate the investigation of the effects of deep-level diversity in social groups within an experimental setting.

Finally, as a social group is formed by dynamic interaction over a longer period of time (Asendorpf & Banse, 2002) and the study of already existing social groups would inevitably lead to a quasi-experimental design, a longitudinal design was chosen for the present study. Out of economical reasons the investigation period was limited to three weeks for each group with four experimental sessions. However, even if systemic empathy probably emerges towards the end of the investigation period, its possible initial levels, as well as the initial levels of individual empathy, should be considered within statistical

analyses. This demands two points of measurement: one measurement of empathy at the beginning and a second measurement of empathy at the end of the investigation period.

The mentioned implications enable the specification of the first hypothesis:

- I. The surface-level diversity of a small group is negatively associated with the level of systemic empathy within this group after the initial period of group formation.*

If the first hypothesis is correct, the level of systemic empathy should be lower in experimental groups with high levels of surface-level diversity, and higher in experimental groups with low levels of surface-level diversity at the end of the investigation period.

Most implications considering the second research questions can be applied for the third research question (small groups, longitudinal design).

As already stated, the experimental setting demands the formation of new social groups. Thus, to investigate the third research question, it is necessary to experimentally implement a shared goal within these social groups. As the perception of a shared goal and not the shared goal itself is the central variable related to relevant group parameters and processes (see Section 2.4.4.2), it is crucial to implement a salient and personally relevant shared goal that all group members feel committed to. The perception of intergroup competition may serve as a factor enhancing the salience of shared goals and goal commitment (Mulvey & Ribbens, 1999).

To achieve the shared goal, the members of the experimental group have to unite in coordination actions. Therefore, the shared goal is hypothesized to elicit mechanisms of cognitive and emotional conversion, leading to higher levels of cognitive and emotional systemic empathy, or to state it more precisely:

- II. The implementation of a shared goal within a small group is positively associated with the level of systemic empathy within this group after the initial period of group formation.*

Again, if the second hypothesis is correct, the level of systemic empathy should be higher in experimental groups with a shared goal, and lower in experimental groups without a shared goal.

Both hypotheses imply, that the two mentioned group parameters are only associated with levels of systemic empathy, not with levels of individual empathy (see Section 2.4). To control this implication, the impact of group parameters on levels of individual empathy will be also tested within this investigation.

To investigate the discussed research questions and hypotheses, a longitudinal experimental design with three experimental conditions was chosen: A control condition, a low-diversity condition, and a shared-goal condition. The methodical details of this design will be presented in the following section.

4. Method

4.1 Participants

4.1.1 Sample Size and Composition

The test persons were recruited by advertisements at several universities (including Freie Universität Berlin, Humboldt Universität Berlin, and Technische Universität Berlin) and by a student mailing list of Freie Universität Berlin. Upon recruitment, interested persons were informed that the present study investigated general aspects of interaction in groups following a longitudinal design and that it will take place at Freie Universität Berlin. The topics of emotions and empathy were not mentioned to prevent any instruction effects on participants' behavior. In a first step, interested persons were asked to report their age and native language – the participants were told that this procedure was necessary to check if they fulfill participation requirements. In fact, this information and the gender of the person were used to assign the participants to the three experimental conditions. According to the research objectives, the groups in the control and shared-goal condition were supposed to be mixed with regard to gender and native language. A further parameter of group composition in these groups was a medium standard deviation in the variable age. In the low-diversity condition, all participants were supposed to have the same sex and the same native language. The standard deviation of the variable age was kept as minimal as possible within this experimental condition.

In sum, 18 groups with 7 participants in each group were formed. The three experimental conditions (control condition, low diversity condition, shared goal condition) comprised 6 groups each.

Persons were only recruited if they had the possibility to attend all four experimental sessions corresponding to their experimental condition. To limit dropout, interested persons were informed about the importance of attending all four experimental sessions and asked to inform the test supervisors as early as possible in case of any hindrances. Participants were offered 60 Euro and the winnings from the Public Goods Game for their participation in the four experimental sessions. Finally, participants were informed about the video supported observation and the anonymity of the gathered data.

However, as several recruited participants did not attend the first experimental session without prior notice, different group sizes were inevitable. The final sample comprised 98 persons in 18 experimental groups with 4-7 participants in each group. Table 2 provides an overview of all 18 experimental groups and the three variables age, gender, and native language that were used for the composition of the 18 experimental groups.

Table 2

Composition of the Experimental Groups at the Beginning of the Investigation Period

Condition	Group	n	Gender Male (%), Female (%)	Age Range, Mean (SD)	Native Language German (%), Other (%)
Control	1	7	2 (28.6), 5 (71.4)	19-41, 31.00 (9.31)	6 (85.7), 1 (14.3)
	2	6	4 (66.7), 2 (33.3)	22-30, 26.50 (3.39)	3 (50), 3 (50)
	3	6	3 (50), 3 (50)	21-37, 28.67 (5.61)	3 (50), 3 (50)
	4	6	3 (50), 3 (50)	24-34, 28.50 (3.51)	4 (66.7), 2 (33.3)
	5	5	1 (20), 4 (80)	19-39, 28.40 (7.54)	3 (60), 2 (40)
	6	4	1 (25), 3 (75)	27-42, 34.00 (6.48)	3 (75), 1 (25)
Low diversity	7	6	0 (0), 6 (100)	24-26, 25.00 (.89)	6 (100), 0 (0)
	8	5	5 (100), 0 (0)	23-26, 24.20 (1.10)	5 (100), 0 (0)
	9	5	0 (0), 5 (100)	22-25, 22.80 (1.3)	5 (100), 0 (0)
	10	4	4 (100), 0 (0)	24-26, 25.00 (.82)	4 (100), 0 (0)
	11	7	0 (0), 7 (100)	25-29, 26.87 (1.35)	7 (100), 0 (0)
	12	4	4 (100), 0 (0)	23-25, 23.75 (.96)	4 (100), 0 (0)
Shared goal	13	6	3 (50), 3 (50)	20-42, 29.83 (7.83)	5 (83.3), 1 (16.7)
	14	6	1 (16.7), 5 (83.3)	28-42, 31.00 (5.40)	5 (83.3), 1 (16.7)
	15	7	2 (28.6), 5 (71.4)	20-31, 26.56 (3.99)	6 (85.7), 1 (14.3)
	16	4	2 (50), 2 (50)	22-28, 24.75 (2.50)	3 (75), 1 (25)
	17	4	1 (25), 3 (75)	20-35, 28.75 (7.09)	3, (75), 1 (25)
	18	6	2 (33.3), 4 (66.7)	22-37, 26.33 (5.61)	4 (66.7), 2 (33.3)

Out of the final sample of 98 persons, 7 participants had to cancel the study before completing all four experimental sessions due to illness and other personal reasons. However, the partial data of these participants was not excluded from the group-level analyses on measures at the beginning of the investigation period. Every group member may represent a part of a specific group dynamic that could be important for systemic empathy, and some analyses require the consideration of all involved participants (e.g., the video supported coding of group discussions; see Schermuly & Scholl, 2012). Thus, group-level analyses on measures at the beginning of the investigation period were computed for 98 participants, and group-level analyses on measures at the end of the investigation period were computed for 91 participants. All individual analyses were computed for 91 participants.

4.1.2 Demographic Characteristics

Table 3 gives an overview of relevant sociodemographic characteristics for the whole sample and separately for the three experimental conditions. Out of economical reasons, the recruitment methods targeted mainly students, thus highly educated (mean years of education = 17.79) and young persons (mean age = 27.43 years) were overrepresented within the total sample. Further, women were slightly overrepresented (61.2 %). Out of this reason, the generalization of the results may be limited, especially on less educated and older populations (see Peterson, 2001, for a meta-analysis on the use of student samples within social science research).

The Levene test for equality of variances (Levene, 1960) revealed that the variable age had a significantly smaller variance within the low diversity condition than within the control condition ($F(1,63)=32.23$; $p<.001$) and than within the shared goal condition ($F(1,62)=20.30$; $p<.001$), indicating a successful experimental manipulation by group composition. The control condition and the shared goal condition did not significantly differ in the variance of the variable age ($F(1,65)=.83$; $p=.37$), which was also in accordance with the research objectives. Further, the percentage of participants with another native language than German did differ significantly between the low diversity condition and the control condition ($\chi^2(1,N=65)=13.42$; $p<.001$), as well as between the low diversity condition and the shared goal condition ($\chi^2(1,N=64)=7.38$; $p=.007$), but not between the control condition and the shared goal condition ($\chi^2(1,N=67)=1.64$; $p=.20$).

Table 3

Demographic Characteristics of the Total Sample and the Three Experimental Conditions

Variables	Total Sample	Control Condition	Low Diversity Condition	Shared Goal Condition
	<i>N</i> = 98	<i>n</i> = 34	<i>n</i> = 31	<i>n</i> = 33
<i>Age</i>				
Mean (SD)	27.43 (5.35)	29.32 (6.31)	24.77 (1.71)	27.97 (5.66)
<i>Sex</i>				
Female (%)	60 (61.2)	20 (58.8)	18 (58.1)	22 (66.7)
Male (%)	38 (38.8)	14 (41.2)	13 (41.9)	11 (33.3)
<i>Education</i>				
Lower secondary education (10 years) ^a (%)	3 (3.1)	1 (2.9)	1 (3.2)	1 (3)
High school (12 years) ^b (%)	57 (58.2)	16 (47.1)	23 (74.2)	18 (54.5)
College / University ^c (%)	38 (38.8)	17 (50)	7 (22.6)	14 (42.5)
<i>Education in Years</i>				
Mean (SD)	17.79 (3.15)	18.37 (3.77)	16.97 (2.32)	17.97 (3.05)
<i>Marital Status</i>				
Single (%)	53 (54.1)	17 (50)	18 (58.1)	18 (54.5)
Long-term partnership (%)	36 (36.7)	10 (29.4)	13 (41.9)	13 (39.4)
Married (%)	7 (7.1)	7 (20.6)	0 (0)	1 (3)
Divorced (%)	1 (1)	0 (0)	0 (0)	0 (0)
Widowed (%)	0 (0)	0 (0)	0 (0)	0 (0)
<i>Nationality</i>				
German (%)	84 (85.7)	27 (79.4)	30 (96.8)	27 (81.8)
Other (%)	14 (14.3)	7 (20.6)	1 (3.2)	6 (18.2)
<i>Native Language</i>				
German (%)	79 (80.6)	22 (64.7)	31 (100)	26 (78.8)
Other (%)	19 (19.4)	12 (35.3)	0 (0)	7 (21.2)
<i>Current Occupation</i>				
Student (%)	70 (71.4)	22 (64.7)	26 (83.9)	22 (66.7)
Full-time employed (%)	3 (3.1)	2 (5.9)	0 (0)	1 (3)
Half-time employed (%)	9 (9.2)	5 (14.7)	0 (0)	4 (12.1)
Housewife / -man (%)	0 (0)	0 (0)	0 (0)	0 (0)
Unemployed (%)	6 (6.1)	2 (5.9)	2 (6.5)	2 (6.1)
Retired (%)	0 (0)	0 (0)	0 (0)	0 (0)
Other (%)	7 (7.1)	2 (5.9)	3 (9.7)	2 (6.1)

Note. ^a German: Mittlere Reife / Realschule. ^b German: (Fach-) Abitur. ^c German: (Fach-) Hochschulstudium, Bachelor / Master

4.2 Procedure

Every participant was supposed to attend four experimental group sessions. The time interval between each session was one week. The initial group composition was not changed within the following experimental sessions; in other words, each participant attended all experimental sessions together with the same group members. All experimental sessions were run by trained research assistants and the author of the study. An experimenter manual containing an overview of measures and experimental applications, standardized instructions, and answers on possible questions by the participants was used to ensure procedure objectivity. To ensure anonymity and data protection within this study, an identification number encoding the gender and the experimental condition was assigned to every participant at the beginning of the first experimental session. All analyses were conducted on basis of the identification numbers and every document linking the identification numbers to the names of the participants was destroyed at the end of the experiment.

4.2.1 First Experimental Session

At the beginning of the first test session all participants were informed about all organizational aspects of the experimental procedure, the anonymity, protection, and scientific purpose of the gathered data, including the video supported observation, and the remuneration for the participation.

Within the shared goal condition, the experimental groups were informed about a task that the group members had to work on together: the creation of a creative collage artistically depicting several emotions. To increase the salience of the shared goal, the participants were told that all groups have to present their collages at the end of the experiment and that it is possible to implement their collage within an exhibition in an associated clinic, if the artwork will be rated positively by independent judges. These instructions were presented orally and additionally in written form, on take-away cards.

The specific shared goal of creating a creative collage was chosen out of several considerations. First, the interrelationship between the difficulty of a shared goal and the two variables goal attraction and goal commitment seems to be u-inverted: Several studies found goal attraction and goal commitment to be high for moderately difficult tasks and low

for easy as well as extremely difficult tasks (e.g., Wright, Contrada, & Patane, 1986). The creation of a collage requires no advanced artistic skill. However, the result can be a valid indicator of the personal level of creativity and intelligence and demands abstract reasoning skills (e.g., Amabile, 1982; Baer, 1996; Ruscio, Whitney, & Amabile, 1998). Thus, the creation of an artistic collage may represent a moderately difficult task. Second, the participants' knowledge that the artwork will be rated and could be exhibited together with the collages of other groups after the experiment was supposed to elicit the perception of an intergroup competitive setting, which is positively associated with goal salience and goal motivation (Mulvey & Ribbens, 1999). Third, it was easy to implement a control intervention within the control and low diversity condition being equivalent to the described shared goal: The participants in these experimental groups were instructed to create individual collages, without interacting with other group members. The creation of an individual collage may be perceived as an individual goal, thus enhancing mechanisms of competition, while a common artwork may enhance mechanisms of cooperation interrelated with systemic empathy (van Mierlo & Kleingeld, 2010). Fourth, the collage could be linked by its topic to the measures and experimental interventions of the four test sessions, which may seem more plausible to the participants than other tasks. And finally, the setting and the materials necessary for this task were less elaborate and expensive than for alternative tasks.

The salience and perception of the shared goal were controlled by two manipulation checks (see Section 4.3.1.7 and Section 4.3.1.8).

After the initial instructions, the members of each experimental group took part in an acquaintance game, which was implemented to enhance group formation processes by social interaction and to reduce anonymity. The short investigation period necessitated such experimental interventions enhancing group identification and formation processes, as the emergence of a shared group-identity is a crucial characteristic of a social group (compare Section 2.4.1) and therefore an important precondition for empirical investigations of social groups.

Following the acquaintance game, the participants completed a demographical questionnaire, individual and systemic measures of empathy, as well as measures of control variables and manipulation checks (see Section 4.3 for details). All measures were preceded by detailed written and standardized oral instructions.

4.2.2 Second Experimental Session

As already stated, the short investigation period necessitated specific methods enhancing group formation processes. Classical group development models as the model by Tuckman (Tuckman, 1965; Tuckman & Jensen, 1977) or Gersick (1988) suggest that systemic socio-emotional processes in groups occur in later stages of the development of a social group, after an initial phase of orientation and group formation. Thus, without further interventions, it may be highly unlikely to evoke and measure systemic empathy in a social group within a period of three weeks.

Out of this reason, an elaborate intervention program was applied in the second and third experimental session to enhance the development of systemic empathy and, by this means, to make differences in systemic empathy that may exist between the experimental conditions measurable. The intervention program was applied identically in all experimental conditions. Thus, differences between the three experimental conditions may not be explained by the intervention program, but at the utmost by interaction effects between the experimental condition and the intervention elements. However, such interaction effects would also be of interest for further practical applications and confirm the research hypotheses from a theoretical point of view.

The intervention methods in the second experimental session targeted the enhancement of cognitive systemic empathy, while the intervention methods in the third experimental session targeted the enhancement of emotional systemic empathy. The intervention methods were designed on the basis of evaluated therapy tools that are commonly used to enhance empathy or associated abilities (e.g., perspective taking).

At the beginning of the second experimental session, all experimental groups participated in a social cognition training that was based on the Movie for the Assessment of Social Cognition (MASC; see Dziobek et al., 2006, for details). The MASC is a movie-based diagnostic tool measuring a person's ability to recognize mental states of other persons, as thoughts, motives, or emotions. This ability is usually referred to as theory of mind, a construct strongly interrelated with cognitive empathy (see section 2.2.1). The MASC is a 15-minute movie about four characters, portrayed by actors. These four fictive persons interact and spend an evening together. The movie is stopped 45 times, and every time the movie is paused, the tested person has to answer a question referring to the thoughts, motives, and

emotions of one of the characters. The MASC has a closed response format with one correct target and three distractor responses.

The MASC was not used as a diagnostic tool within this study but only for intervention purposes. The items of the MASC were extracted from the test and the film-based stimuli were presented to each experimental group. After presenting the item, each group had to discuss, which of the four presented possible responses was correct. Subsequently, the participants had to choose one of the responses and provide the answer as a group. This was followed by an immediate feedback of the experimenter, telling the participants if their response was correct or not, and why. The feedback on the reasons for the correctness or incorrectness of an answer encompassed elaborate psychoeducative input, e.g. on specific gestures or facial expressions that are associated with certain emotions.

In this form, this newly developed intervention element was comparable to common video-based therapy tools aiming at enhancing theory of mind, cognitive empathy, and perspective taking abilities (e.g., Charlop-Christy & Daneshvar, 2003; Delano, 2007). However, comparable methods were not applied within a group setting before. The interaction between the group members was hypothesized to enhance the understanding of how the other group members understand the emotions of other persons – this form of group-specific, meta-level cognitive empathy may lead to higher systemic cognitive empathy.

As a second element of the second experimental session participants were trained in decoding the facial expressions of other group members. Common psychological intervention programs that are used to enhance emotion recognition abilities usually are based on stimuli as photographs of facial expressions (e.g., Elfenbein, 2006). Participants have to guess what emotion is depicted on the photograph, based on a multiple-choice response format with one target response and several distractors. Afterwards they are informed if their response was correct or not, and why. Some intervention programs are using more naturalistic and therefore more ecologically valid settings as face-to-face interactions or video modelling to train emotion recognition abilities (Constanzo, 1992; Feldman, Philippot, & Custrini, 1991).

The group context of the present investigation called for the development of a similar intervention program, focusing on a *systemic* training of cognitive empathy. For this

purpose, the members of each experimental group were divided in two teams and played a game that was similar to the party-game TABU (Hersch, 1990). One player drew a card from a deck, which was not visible to other persons. The card listed an emotion word, e.g. “*angry*” or “*surprised*”. Now the player had to depict the listed emotion by facial expressions, without using words, while his or her team members had to guess the correct emotion. If the emotion was guessed correctly, the team was given a point and the participant had the possibility to draw another card and to depict another emotion, until he or she exceeded a time limit of 60 seconds. More correctly guessed emotions within the time limit were rewarded with more points. After time was over, the emotions that were not recognized were revealed to all group members. Several studies demonstrated that this kind of feedback is strongly associated with the success of intervention programs aiming to enhance the ability to recognize the emotions of another person (Beck & Feldman, 1989; Elfenbein, 2006; Gillis, Bernieri, & Wooten, 1995). Subsequently, a member from another team took over and the procedure was repeated. In sum, this intervention element was limited to 30 minutes, or approximately 20-25 rounds.

At the end of the second experimental session participants in the shared goal condition were given 30 minutes to work on their common artistic collage, while participants in the control and low diversity condition were instructed to work on the individual artworks without consulting the other group members.

Afterwards, participants had to complete several short questionnaires measuring control variables and the effectiveness of the described experimental manipulations (see Section 4.3.1 for details). After completing the questionnaires, participants were dismissed.

4.2.3 Third Experimental Session

The intervention methods within the third experimental session were designed to enhance systemic emotional empathy. Usually, participants of emotionally oriented empathy training programs have to deal in detail with own emotions and the emotions of another person, often within an interaction or conversation with the target person (e.g., Erera, 1997; Pecukonis, 1997). However, emotionally oriented empathy training programs are usually developed for a very specific purpose and within a limited context, e.g., for the treatment of sex offenders (Wastell, Cairns, & Haywood, 2009), elementary schoolchildren (Sherman, 2008), hospital staff nurses (Herbek & Yammarino, 1990), etc. Thus, most

intervention programs addressing emotional empathy may be strongly limited with regard to their generalizability.

Further, the effectiveness of these interventions seems to be lower than the effectiveness of cognitively oriented intervention programs (Erera, 1997). Besides of methodological shortcomings, an important reason for the ineffectiveness of these interventions may be the persistent lack of proven guidelines on the use of appropriate and reliable methods for the induction of specific emotions within an experimental setting (Rottenberg, Ray, & Gross, 2007). As emotional empathy involves sharing the emotions of another person, an emergence of congruent emotional states is a crucial condition for the success of any intervention program addressing emotional empathy.

Thus, the intervention elements within the third experimental sessions combined empirically proven methods of emotion induction with classical elements of emotionally oriented empathy trainings, as conversations on personal emotional states, adapting these classical elements for the systemic context of the present investigation. All induction methods were limited to the induction of pleasant emotions due to ethical considerations.

Several researchers identified films as the most effective method to elicit specific emotional states in a laboratory setting (e.g., Rottenberg et al., 2007; Westermann, Spies, Stahl, & Hesse, 1996). Within these and many other experimental studies, short excerpts of popular movies have been successfully used to elicit specific and distinct emotions, e.g., Jurasova and Spajdel (2013) successfully used a scene from the movie "When Harry met Sally" (Reiner, 1989) to elicit distinct amusement, a scene from the movie "Schindler's List" (Spielberg, 1993) to elicit distinct sadness, etc. Gross and Levenson (1995) demonstrated that expert validation by emotion researchers, film scholars, or similar experts is a reliable method to identify adequate film excerpts from commercial films, which may be used for the induction of specific and distinct emotions within a laboratory setting.

Based on these findings and considerations, an excerpt from the movie "The Pursuit of Happiness" (Muccino, 2006) was used at the beginning of the third experimental session to induce the distinct basic emotion happiness. Specifically, the last 12 minutes of the movie were shown to all participants, where the protagonist of the movie (Chris Gardner, portrayed by Will Smith) wins a highly lucrative and coveted full-time position in a stock brokerage firm. Similarly as in the study by Gross and Levenson (1995), several emotion researchers and film scholars from the cluster "Languages of Emotion" at Freie Universität

Berlin validated the chosen film excerpt with regard to its effectiveness to elicit happiness as a distinct emotion.

The emotion happiness was selected because it is regarded as a basic emotion (Ekman, 1992) and may therefore be similarly experienced by all group members, independently from their cultural background. Further, according to the circumplex model of affect (Russell, 1980; Russell, Lewicka, & Niit, 1989), happiness is the only basic emotion that is cross-culturally and consistently experienced as pleasant. Thus, the selection of happiness met the already mentioned criterion of inducing only pleasant emotions.

The presentation of the film excerpt was combined with the task to identify the emotions of the protagonist. After seeing the film excerpt, the recognized emotions were discussed within the group. This additional training tool was in line with the intervention methods of the second experimental session.

After watching the film excerpt, all participants were instructed to remember five personal positive life events, which have been emotionally important, intense, but not intimate for them. Intimate life events were excluded out of ethical considerations. Accordingly, participants were instructed to remember the emotions that they experienced during these life events. This reminiscence technique is recognized as a further effective method to elicit specific emotional states (Westermann et al., 1996) and served as a preparation for the following dyadic interaction task.

After completing the reminiscence exercise the participants were told to describe the remembered life events and the experienced emotions as precisely as possible to all other group members within dyadic interactions. This intervention element is not only in line with already mentioned emotional empathy training programs (Erera, 1997; Pecuconis, 1997). Further, such dyadic interaction tasks may also serve as another effective method to elicit specific emotional states (Roberts, Tsai, & Coan, 2007), which is an important condition for the experience of emotional empathy. This intervention element was limited to 40 minutes and was therefore the main training tool within the third experimental session.

After completing the dyadic interaction task, participants in the shared goal condition were again given 30 minutes to work on their common collage, and participants in the control and low diversity condition were given 30 minutes to work on their individual collages. Finally, participants again had to complete questionnaires measuring control

variables and the effectiveness of the described experimental manipulations (see Section 4.3.1 for details). After completing the questionnaires, participants were dismissed.

4.2.4 Fourth Experimental Session

At the beginning of the last experimental session, the participants in the shared goal condition had to present and explain their artistic collage as a group and received feedback on their work. Afterwards, participants completed the same measures as in the first experimental session, including identical instructions and procedures. Additionally, participants completed two personality measures that were included as control variables (see Section 4.3 for details).

At the end of the experiment, all participants were debriefed about all details of the investigation and kindly asked not to inform other participants or persons interested in participation about the theoretical background of the study.

The experimental procedure within all four experimental sessions is summarized in Figure 6.

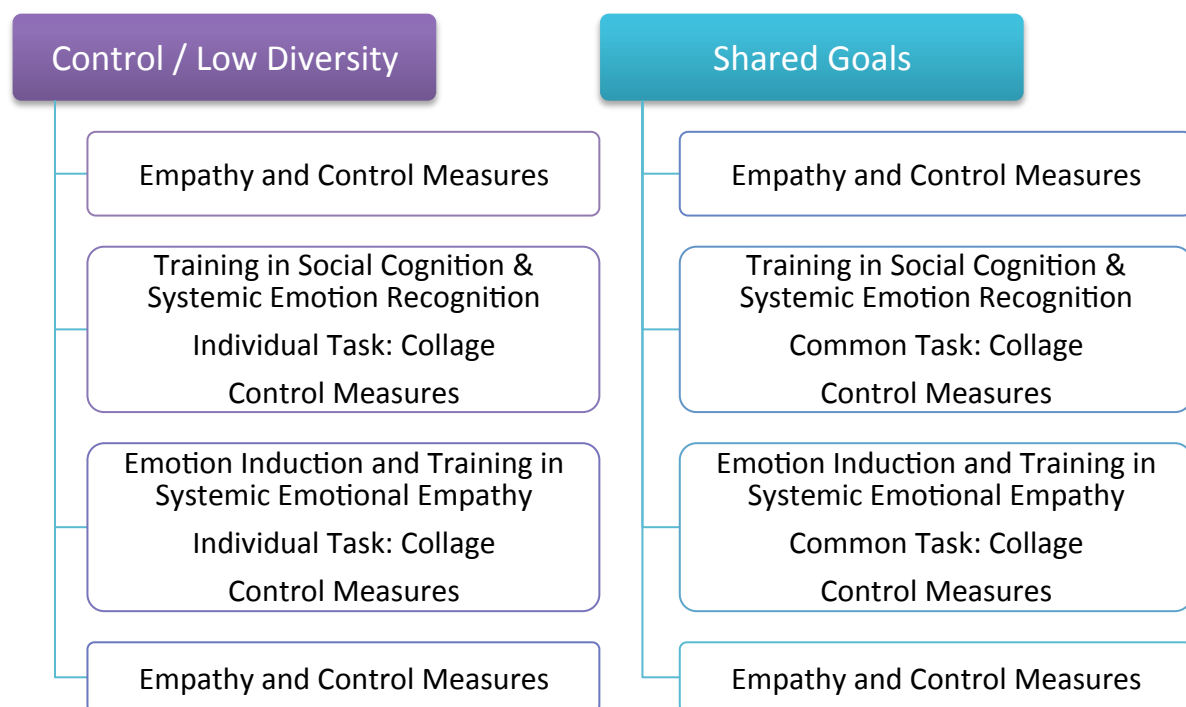


Figure 6. The Experimental Procedure of the Present Study

4.3 Measures

The measures of the present investigation comprised several questionnaires on individual and systemic empathy and related individual and systemic control variables. Further, measures comprised the Multifaceted Empathy Test (MET; Dziobek et al., 2008), an image-based measure of individual empathy, a Public Goods Game as an indirect measure of systemic empathy, and group discussion parameters representing systemic levels of empathic communication as well as related systemic control variables. The measurement of individual and systemic empathy with several different methods follows the multitrait-multimethod approach (Campbell & Fiske, 1959; see Section 3).

4.3.1 Questionnaires

4.3.1.1 IRI

Participants completed the Interpersonal Reactivity Index (IRI; Davis, 1983) as a measure of individual empathy in the first and fourth experimental session. The IRI is based on a multidimensional operationalization of empathy (see Section 2.1.2) and includes four scales: Perspective Taking and Fantasy, as subdimensions of cognitive empathy, and Empathic Concern and Personal Distress, as subdimensions of emotional empathy.

For the present investigation, the German adaptation of the IRI by Paulus (2009) was applied. This shortened version contains 16 items with 4 items for each of the four scales. Respondents rated items as “Before criticizing somebody, I try to imagine how I would feel if I were in their place” (for Perspective Taking) using a 5-point agreement response format ranging from 1=“never” to 5=“always”.

The IRI by Davis (1983) is one of the most common conceptualizations of empathy and highly reliable: Davis (1983) reports a Cronbach’s Alpha ranging from .71 to .78 for the four scales and a re-test reliability ranging from .62 to .80 for an interval of 8 to 10 weeks. Paulus (2009) reports even higher values. Further, Davis (1980; 1983; 1994) reports high convergent and divergent validity of the IRI, as well as associations with several external criteria, such as social functioning, self-esteem, or emotionality. Paulus (2009) replicated the high convergent, divergent, and external validity indices of the IRI.

However, the IRI also suffers from the limitations already discussed in Section 2.1.2 and in Section 3: From a conceptual point of view, the IRI poorly differentiates between

empathy and related constructs as perspective taking or compassion and does not consider several basic functions that are associated with cognitive and emotional empathy. Moreover, the IRI is limited due to the disadvantages of self-report measures, e.g. biases because of cognitive distortions, response-tendencies, and social desirability. Thus, the MET (Dziobek et al., 2008) was used as an additional image-based measure of individual empathy (see Section 4.3.2).

4.3.1.2 CEEQ-Group

Participants completed an adaptation of the Cognitive and Emotional Empathy Questionnaire (CEEQ; Savage, Dziobek, Teague, & Borod, submitted) as a self-report measure of systemic empathy in the first and fourth experimental session.

The CEEQ is a newly developed measure of cognitive and emotional empathy and based on a multi-dimensional conceptualization of empathy being similar to the conceptualization of the IRI (Davis, 1983). It consists of 30 items that are assigned to four scales: Mental State Perception and Perspective Taking as subscales of cognitive empathy, and Mirroring and Empathic Concern as subscales of emotional empathy. Test persons have to rate statements as “When I hear that other persons laugh, I feel the need to laugh, too” (for Mirroring) using a 5-point agreement response format ranging from 1=“not at all true” to 5=“very true”. Savage et al. (submitted) report an adequate to good reliability (Cronbach’s Alpha ranging from .52 to .84), as well as high convergent, divergent, and external validity of the four subscales.

The CEEQ by Savage et al. (submitted) represents a measure of individual empathy. For the present investigation, the original items of the CEEQ were adapted into systemic measurements by targeting the members of the experimental group that the test person was part of. For example, a statement of the original CEEQ as “When I hear that other persons laugh, I feel the need to laugh, too” was changed into “When I hear that other persons in our group laugh, I feel the need to laugh, too”. Written and oral instructions were used to ensure that the test person answered the items with regard to his or her experimental group within the present study and no other group. The original response format of the CEEQ was not changed.

Considering the test properties of the original CEEQ (Savage et al., submitted), the CEEQ-group may represent a reliable and valid systemic self-report measure of empathy. In

contrast to the IRI, the CEEQ differentiates between some constructs related to empathy, as emotional contagion (Savage et al., submitted). However, it is also affected by the general disadvantages of self-report measures and was therefore supplemented by alternative systemic measures of empathy, the Public Goods Game (see Section 4.3.3) and group discussion parameters (see Section 4.3.4).

4.3.1.3 SDS

As empathy is generally regarded as socially desirable (Cohen & Strayer, 1996) and as several studies have found associations between self-report measures of empathy and indexes of social desirability (Eisenberg & Fabes, 1990), it is crucial to include the individual tendency to provide socially desirable responses as a control variable when using self-report measures of empathy. Thus, participants completed a German version of the Social Desirability Scale (SDS; Stöber, 1999; 2001) in the fourth experimental session to control for the influence of social desirability on the main variables.

The SDS consists of 17 items measuring a person's tendency to give socially desirable responses, e.g. "I always stay friendly and courteous with other people, even when I am stressed out". Test persons rated the items following a true (=1) / false (=0) response format.

Stöber (2001) reports the SDS to be a highly reliable scale (Cronbach's Alpha = .72; four-week retest-correlation of $r=.82$) with high convergent and discriminant validity.

However, because changes in social desirability across the test interval were not of interest, and considering the high retest-reliability, the SDS was administered only once in the fourth experimental session.

4.3.1.4 NEO-FFI

As some authors report associations between empathy and the Big Five personality traits, especially Agreeableness and Extraversion (e.g., van der Zee, Thijs, & Schakel, 2002), it is possible that these personality factors influence the development of individual and systemic empathy across the investigation period.

Accordingly, a German version of the NEO-Five-Factor Inventory (NEO-FFI; Borkenau & Ostendorf, 1993; see Costa & McCrae, 1992, for the original version) was administered in the fourth experimental session to control for associations between the main variables of the present study and the Big Five personality traits: Neuroticism, Extraversion, Openness,

Agreeableness, and Conscientiousness (see McCrae & Costa, 1987, for a detailed description and validation of the Big Five personality model).

The NEO-FFI is a 60-item questionnaire with 12 items for each of the Big Five personality traits. Test persons rate items as “I often feel tense and jittery” (for Neuroticism) following a 5-point agreement response format ranging from 1=“totally disagree” to 5=“totally agree”. The NEO-FFI is one of the most common psychological self-report measures with excellent reliability and validity indices (see Borkenau & Ostendorf, 1993, for details).

4.3.1.5 Questionnaire on Group Processes

As empathy within groups can highly depend on the level of identification of the group members with their group (e.g., in case of empathy-motivated helping behavior; see Simon, Stürmer, & Steffens, 2000), as well as on group climate (e.g., Johnson, Burlingame, Olsen, Davis, & Gleave, 2005), a questionnaire on group processes was administered to control for these variables. As group identification and group climate are associated with cognitive and affective states (Carr, Schmidt, Ford, & DeShon, 2003), as well as with changeable group dynamics (Choi, Price, & Vinokur, 2003), both may represent rather systemic states than traits with reduced stability across time and a low cross-situational consistency. Thus, the questionnaire on group processes was administered after every experimental session.

The questionnaire consists of four items measuring group identification that were based on a scale by van Dick, Knippenberg, Hägele, Guillaume, and Brodbeck (2008) and three items measuring group climate that were based on a scale by Aubé and Rousseau (2005). Participants rated items as “I define myself as a member of my group” (for Group Identification) based on a 5-point agreement response format ranging from 1=“totally disagree” to 5=“totally agree”.

4.3.1.6 Questionnaire on Group Composition

Theoretical considerations and several empirical investigations speak for a possible association between *perceived* surface-level diversity and systemic empathy (see Section 2.4.4.1). Thus, to investigate this association, it is not sufficient to successfully manipulate the surface-level diversity of experimental groups in one experimental condition. In addition,

the theoretical and empirical background presented in Section 2.4.4.1 points to the importance of a crucial manipulation check: to control for the participants' *perceptions* of surface-level diversity. More specifically, it is necessary to control if participants' perceptions of surface-level diversity were lower in the low-diversity condition than in the control and shared-goal condition. For this means, a questionnaire on group composition was administered in the first and the fourth experimental session.

The questionnaire was based on a scale by van Dick et al. (2008). One item was used to measure the general level of perceived surface-level diversity ("How similar or different are the members of your group in general?") and three items were used to measure the perceived surface-level diversity in the three main variables gender, age, and ethnic background (e.g., "How similar or different are the members of your group with respect to their age?"). As van Dick et al. (2008) demonstrated that the effects of perceived surface-level diversity on group identification and systemic emotions are moderated by diversity beliefs, or in other words are dependent on whether a person *believes* that diversity is beneficial for a group or not, four additional items on diversity beliefs were added to this questionnaire (e.g., "Heterogeneity is of great use within a project team"). Participants rated perceived diversity on a 5-point likert scale ranging from 1="very similar" to 5="very different" and diversity beliefs based on a 5-point agreement response format.

4.3.1.7 Questionnaire on Goal Representation

A goal can be defined as an internal representation of a desired state (compare Section 2.4.4.2). Thus, a successful experimental implementation of a shared goal requires participants' perception and representation of this goal as well as goal commitment. To control if the experimental manipulation in the shared goal condition (see Section 4.2.1) elicited a shared representation of this goal and goal commitment, participants in the three experimental conditions completed a questionnaire on goal representation in the first and fourth experimental session.

The questionnaire contained eight items. The first item asked for a general perception of a shared goal ("Our group has a shared goal") with a closed response format ranging from 1="totally disagree" to 5="totally agree" and was supplemented by an item with an open response format to check if the represented goal was the collage that the participants had to work on ("If you rather agree on the first question, please specify which

goal your group pursues. If it is more than one goal, please specify all of them"). This control measure was based on the manipulation check used by van Mierlo and Kleingeld (2010).

Six additional items were used as a measure of goal commitment. The items were based on a scale by Aubé and Rousseau (2005), which is in turn an adaptation of a measure by Klein, Wesson, Hollenbeck, Wright, and DeShon (2001). Participants rated items as "I am dedicated to pursue the shared goal" following a 5-point agreement response format ranging from 1="totally disagree" to 5="totally agree".

4.3.1.8 Questionnaire on Task Motivation and Orientation (Shared Goal Condition)

To monitor if the participants in the shared goal condition were motivated to pursue the shared goal, which may be an important condition for the success of the experimental manipulation (see Section 2.4.4.2), a measure of task motivation was administered in the second, third, and fourth experimental session. Further, to make sure if the experimental manipulation described in Section 4.3.1 elicited cooperation and not competition in the shared goal condition, which is an important difference between shared and individual goals (compare Section 2.4.4.2), a scale on cooperative and competitive orientation during interaction was added to this questionnaire. This questionnaire was not administered in the first session, because the items were related to the teamwork on the collage, which started in the second experimental session.

The questionnaire entailed 12 items. Three items were used to measure intrinsic motivation and three items were used to measure extrinsic motivation. These first six items were derived from Cooper and Jayatilaka (2006). Further, cooperative and competitive orientation during teamwork was measured by three items each, based on a scale by van Mierlo and Kleingeld (2010).

Participants rated items as "When working on the task, I wanted to be better than the other group members" (for Competitive Orientation) based on a 5-point agreement response format ranging from 1="totally disagree" to 5="totally agree".

4.3.2 Multifaceted Empathy Test (MET)

The Multifaceted Empathy Test (MET; Dziobek et al., 2008) served as an image-based measure of individual empathy and was administered in the first and fourth experimental session. The MET allows a separate measurement of cognitive and emotional empathy and is

based on images depicting human emotional expressions from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008). In case of cognitive empathy, respondents have to recognize the correct emotion that is depicted on the image with one correct target and three false distractor responses. In case of emotional empathy, respondents have to indicate how much they feel with the person who is depicted, based on a 9-point likert scale. For the present study, a shortened version of the MET with 30 items for cognitive empathy and 30 items for emotional empathy was used. For further details of test construction and administration, see Dziobek et al. (2008).

The MET is based on a multidimensional conceptualisation of empathy, which is similar to the operationalization of empathy within the IRI (Davis, 1983). However, by using images the MET can be considered as being more naturalistic and therefore more ecologically valid than conventional self-report measures of empathy (Dziobek et al., 2008; Dziobek, 2012). Further, Dziobek et al. (2008) report a high reliability as well as a high convergent and discriminant validity of the test.

Disadvantages of the MET can be seen in an insufficient differentiation between empathy and related constructs (e.g., emotional contagion or perspective taking), as well as in its vulnerability to several types of response biases with regard to the items on emotional empathy (e.g., social desirability). Moreover, even if the MET is more ecologically valid than conventional self-report measures, it may be questionable to which extent its results correspond with a person's actual thoughts and behavior (Dziobek et al., 2008).

Thus, the MET may serve as an additional measure of individual empathy, but not as the exclusive operationalization of individual empathy within this study.

4.3.3 Public Goods Game

Empathy is substantially associated with prosocial behavior in general and specifically with altruism (see Section 2.1.3) as well as with cooperation (see Section 2.3.1). Thus, implementing an additional systemic operationalization, which directly measures these most common behavioral outcomes of empathy, may result in a higher incremental validity of the present investigation.

For this means, a public goods game was administered as a measure of systemic empathy in the first and fourth experimental session. The design of the game was based on the procedure used by Fehr and Gächter (2000; 2002): All test persons received a small

amount of money (4 Euro, in 20 x 20-Cent coins). Every test person had the possibility to decide anonymously how much of this money he or she kept for him- or herself, and how much of it he or she paid into a common cash box. All money in this common cash box was multiplied by 1.5 at the end of the game. The final amount of money in the common cash box was divided into as many equal parts as the group had members, and paid to them (e.g., when there were 12 Euro in the common cash box of a group with six members, 12 was multiplied by 1.5 = 18, 18 was divided by 6, and 3 Euro were paid to every group member).

Thus, if a group as a system cooperates perfectly, every group member would pay all the money into the common cash box, so that all members of the group would profit. However, there is also the possibility to be a *freerider*: if a person keeps the money for him- or herself and hopes to profit from the generosity and cooperativeness of the other group members, such a person would achieve the maximal individual profit. According to the upper example, if a person in a group of six persons keeps the 4 Euro for him- or herself, and the other group members pay 12 Euro into the common cash box, the *freerider* would receive $4 + 3 = 7$ Euro in total within the game. A person of the same group that paid all his or her money into the common cash box would receive only 3 Euro.

In other words, in a public goods game there is a difference between a maximal systemic profit and a maximal individual profit. Pursuing a systemic profit rather than an individual profit may represent a prosocial orientation and a high level of cooperativeness, while pursuing an individual profit on cost of others rather than a systemic profit may represent an antisocial orientation and a low level of cooperativeness. Subsequently, the amount of money that was paid into the common cash box served as a systemic measure of prosocial orientation and cooperativeness, and therefore as an indirect measure of systemic empathy.

For further details on the design, administration, and outcomes of public goods experiments see Fehr and Gächter (2000; 2002) and Holt and Laury (1997).

4.3.4 Group Discussion Parameters

Empathy is associated with several patterns of interactive behavior, such as empathic listening, openness, paraphrasing, and reflection (Martinovski, Traum, & Marsella, 2007). Davis (1994) describes a positive association between empathy and the general quality as well as the quantity of face-to-face communication in general. Furthermore, some authors

conceptualize empathy and related constructs, e.g. emotional mimicry, as a communicative act: they postulate that the evolutionary functions of empathy as cooperation and prosocial behavior (see Section 2.1.3 and 2.3.1) are mainly realized by verbal and nonverbal communication (Bavelas, Black, Lemery, & Mullett, 1986; Gladstein, 1983; Han, 2008).

Thus, parameters of empathic communication were included as a further measure of systemic empathy in the present investigation. Communication parameters were derived from the video supported observation of a group discussion, which was administered in the first and fourth experimental session. To ensure that the interactions during the group discussion were as independent from the experimenters as possible, all experimenters left the laboratory during the group discussion. To ensure procedure objectivity, the topic of the discussion was standardized across all experimental conditions, as well as the preceding instructions. The selection of the topic of the group discussion was based on several criteria.

First, to make the measurement of empathic communication possible at all, the topic of the discussion had to be controversial enough to elicit socioemotional reactions and statements from as many group members as possible. For this reason, the topic should be part of a recent public controversy with popular pro and contra arguments. This excluded all topics in which one position is widely accepted within German Society, e.g. if German companies and institutions should provide more family-friendly workplaces (88% of all Germans agree to this statement, see a study by the Federal Ministry of Family, Senior Citizens, Women, and Youth, 2005).

Second, due to ethical and methodological considerations, the topic should not elicit too intense emotional reactions, which may lead to intense conflict and unpleasant emotional states. Besides the moral issues related to such an experimental intervention, strong conflicts within an experimental group may lead to a disturbed group dynamic, which may affect subsequent measures as well as the interventions that were implemented in the second and third experimental session (see Section 4.2.2). This excluded all highly controversial topics that are usually associated with intense emotional reactions and central personal values or beliefs, e.g. abortion (see Jelen & Wilcox, 2003, for a review on public opinion on abortion and associated behaviors).

Third, forming an opinion on the selected topic should not require advanced technical or subject-specific knowledge, as more persons respond emotionally to such “easy” issues than to “hard” issues (Carmines & Stimson, 1980).

Based on these criteria, the question if smoking should be banned completely in all public places in Germany was chosen as the topic of the group discussion in all experimental conditions. Public opinion research suggests that this topic is discussed rather controversially in Germany (35% of all Germans argue for a total ban of smoking in public places, 14% argue for no ban at all, and 48% argue for a compromise solution with separate smoking areas, see Institute for Public Opinion Research Allensbach, 2008). However, as attitudes towards smoking are highly complex, context-dependent, and partially implicit (Sherman, Rose, Koch, Presson, & Chassin, 2003), emotional reactions during a debate on a smoking ban may be not strong enough to elicit intense conflicts within a group. Finally, having an opinion on a smoking ban does not require any advanced technical knowledge.

Participants were instructed to discuss freely and informally on the topic. The whole discussion was recorded on two camcorders, ensuring a high resolution (1920x1080), a good visibility of the body and the face of each participant, and a good audio quality.

The recorded group discussions were analyzed by two trained research assistants. Communication parameters were derived from a standardized coding of each verbal and nonverbal interaction element within the group discussion. The coding was based on the Discussion Coding System (DCS; Schermuly & Scholl, 2012; Schermuly, Schröder, Nachtwei, & Scholl, 2010). The DCS is a recently developed instrument, which enables a profound analysis of verbal and nonverbal communication among humans by coding functional as well as interpersonal meanings that are associated with each communicative act. Central interpersonal variables that are coded within the DCS are Dominance and Affiliation, as two central parameters of human communication in general (Scholl, 2013; Luxen, 2005). Therefore, the average levels of dominance and affiliation were included as systemic control variables within the present investigation.

Further, the DCS measures if a communicative act entails negative or positive socioemotional expressions, a question, a proposal, an agreement, or a rejection. By clearly distinguishing the sender and receiver(s) of each communicative act it is partially possible to analyse the relationships between the interacting persons and not only individual communication parameters. Thus, the DCS can be regarded as a systemic measure. The DCS is a highly economic, reliable, and valid instrument, especially with regard to interrater-agreement and ecological validity (Schermuly & Scholl, 2012; Schermuly et al., 2010).

To measure parameters of empathic communication within the group discussions, an adaptation of the DCS by Wacker (2011) was administered. Wacker (2011) operationalized empathic communication by coding the level of cognitive and emotional empathic expressions for each communicative act (see Section 8.1 for details).

For the sender, levels of verbal empathic communication ranged from 0="no empathic communication" over 1="repetition" (e.g., paraphrasing of emotional statements), and 2="elaboration/evaluation" (e.g., inquiry on the other's emotions), up to 3="explicit empathic reaction" (e.g. expressing an understanding for the other's emotions). The sender's level of empathic communication was coded for cognitive and emotional empathy separately. For the receiver(s), several nonverbal, paraverbal, verbal reactions were coded as empathic feedback (e.g., validating expressions as "sure!").

Summing up, the operationalization of empathic communication by Wacker (2011) differentiates between sender vs. receiver, cognitive vs. emotional empathy, verbal vs. nonverbal empathic communication, and additionally codes the level of empathic communication on a scale ranging from 0="no empathic communication" to 3="explicit empathic reaction" for the sender.

Wacker (2011) reports an acceptable interrater-reliability for her adaptation of the DCS (Schermuly & Scholl, 2012) ranging from $ICC=.66$ for the sender's level of verbal emotional empathic communication, and $ICC=.70$ for the sender's level of verbal cognitive empathic communication, up to $ICC=.77$ for the receiver's verbal and nonverbal empathic feedback. Further, Wacker (2011) reports a high convergent and divergent validity of the adaptation.

4.4 Statistical Analyses

All data were entered into the program "Statistical Package for the Social Sciences" (SPSS, version 22.0). Two research assistants independently entered 15% of the data, which were randomly selected, twice to control for the accuracy of data entry. By comparing the data of the two independent assistants a rate of errors per entrances was computed. The error rate was less than 1 per 1000 entrances ($0.7 / 1000$), thus the accuracy of the data entry was considered as acceptable. Hence, only one research assistant entered the remaining 85% of the data. Reversed items were recoded.

Before analyzing the data, data were prepared by replacing missing values and adjusting outliers. In a first step of analysis, descriptive statistics and reliabilities were computed for all measures of the present investigation. Second, the effectiveness of the experimental manipulation was tested by analyzing the measures of perceived diversity (see Section 4.3.1.6), goal representation (see Section 4.3.1.7), and task motivation (see Section 4.3.1.8). Third, associations between empathy measures and demographics as well as further personality measures were analyzed to determine relevant control variables for the main analyses. Fourth, to answer the first research question, partial correlations between individual and systemic measures of empathy and confirmatory factor analyses with the latent higher-order factors individual and systemic empathy were computed. Fifth, analyses of covariance for repeated measures (ANCOVAs) were computed to test for the effect of diversity and a shared goal on the development of individual and systemic empathy over the investigation period, and multilevel analyses were additionally computed to test for the effect of diversity and a shared goal on the development of systemic empathy over the investigation period. Finally, multiple regressions were administered as follow-up analyses to control for moderating and mediating effects of control variables.

The details of the statistical analyses will be presented in the following Sections.

4.4.1 Data Preparation

Treatment of Missing Data

As SPSS does not allow the estimation of parameters from incomplete data, missing values of a participant lead to the discard of the entire unit within each analysis. This results in an inevitable loss of power (Schafer & Graham, 2002), which is especially disadvantageous in studies with small sample sizes, as in the present investigation. To prevent additional loss of power and preserve the data from as many participants as possible, missing data were imputed in the present study. As mean imputation, the most common imputation method, may lead to biased data and has several severe disadvantages (Little & Rubin, 2002), the “*Expectation-Maximization*” (EM) algorithm was applied to impute missing data. If the MCAR-condition (*missing completely at random*) is met within a dataset with missing values, the EM algorithm provides valid estimates for missing mean values, standard deviations, and covariances (Allison, 2009). To test if the missing data of the present study met the necessary MCAR-condition, Little’s MCAR-test (Little, 1988) was applied for all subscales. The

null hypothesis of missing completely at random was not rejected for the analyzed data set ($\chi^2(1,612)=534.31; p=.99$).

Therefore, missing data were replaced by EM imputation at the level of subscales in the present investigation, with one exception: missing data from the seven participants, who did not attend the fourth experimental session (see Section 4.1.1), were not replaced by mean imputation, as these missing values were missing systematically and therefore not spread uniformly across the data matrix. In such a case, EM imputation may lead to biased estimates (Schafer & Graham, 2002; Allison, 2009). In sum, 115 missing values were replaced by EM imputation at the level of subscales, with 5880 subscale values in total (ratio: 1.96%).

Treatment of Outliers

The most common practice to detect univariate outliers in data samples is the *three-standard-deviations rule*: All values that are more than three standard deviations above or below the statistical mean are considered as outliers and either discarded or replaced with a value exactly three standard deviations above or below the mean. This practice has several disadvantages (Leys, Ley, Klein, Bernard, & Licata, 2013): First, the statistical mean and standard deviations are strongly affected by outliers, or, simply put, both indicators are already part of the problem and therefore they cannot serve as valid indicators to detect the problem. Second, the *three-standard-deviations rule* is based on the assumption that the analyzed distribution is normal: Within a normal distribution, any value more than three standard deviations below or above the mean is in the lower or upper 1% of the distribution, thus it is likely that this value is discordant. However, outliers heavily affect the distribution and often prevent it to be normal. Thus, a model that is based on normal distribution to detect outliers, while outliers heavily affect each normal distribution, is also already part of the problem and cannot be regarded as an appropriate procedure to solve it. Third, the critical distance from the statistical mean which is used to define a value as discordant highly depends on sample size: The criterion of three standard deviations results in a high probability to detect outliers in bigger sample sizes, but with a low probability to detect outliers in smaller sample sizes (see Barnett & Lewis, 1994, p.223).

Thus, the median absolute deviation (MAD), which was proposed by Leys et al. (2013) as a more robust alternative to the *three-standard-deviations rule*, served as a criterion to detect univariate outliers in the present investigation. Values with more than three median

absolute deviations were considered as outliers. To preserve power, outliers were not discarded but replaced with values corresponding to the median plus or minus three times the MAD. 54 out of 5880 subscale values (0.92%) were identified as outliers and replaced.

Multivariate outliers were inspected in all analyses of covariance for repeated measures (Section 5.5) and all regression analyses (Section 5.6) by determining the Mahalanobis distance of each case (Tabachnik & Fidell, 2007). Cases with a significant Mahalanobis distance ($p < .001$) indicating the case to be a multivariate outlier were not detected in the present sample. The absence of multivariate outliers may be a result of replacing univariate outliers (Tabachnik & Fidell, 2007).

LN-Transformation

A Kolmogorov-Smirnov test for normality (Massey, 1951) was applied to determine if subscales of the prepared data sample were normally distributed, as normal distribution is an important condition for almost all statistical analyses in the present study. Out of 64 subscales, the distribution of 47 subscales did significantly differ from normal distribution ($p < .05$). As a consequence, all subscales were log-transformed with $\log(x+1)$, based on the *Box-Cox transformation technique* (Box & Cox, 1964; see also Sakia, 1992, for a review).

For reasons of clarity and comprehensibility reports of descriptive statistics in Section 5.1 are based on the original scores. All following analyses as reliabilities, correlations, confirmatory factor analyses, analyses of covariance, multilevel analyses of covariance, and regressions are based on log-transformed scores.

4.4.2 Preliminary Analyses

Basic descriptive statistics as mean, range, and standard deviation were computed for the subscales of all measures to provide a general overview. Descriptive statistics were computed for the total sample and separately for the three experimental conditions. If a univariate analysis of variance revealed that the mean of a control measure significantly differed between the three experimental conditions ($p < .05$), this subscale was included automatically as a covariate in the main analyses. Further, reliabilities of all measures were computed, both in terms of internal consistencies and, if the measure was administered twice or more often, in terms of re-test reliabilities.

To test if the experimental manipulation with regard to group composition was successful and if it led to a lower perceived surface-level diversity in the low diversity condition, analyses of variance with experimental condition as independent variable and the measure of perceived surface-level diversity as dependent variable (see Section 4.3.1.6) were computed. Analyses of variance were computed for all items of the subscale Perceived Diversity, and separately for general surface-level diversity and surface-level diversity in the three variables age, gender, and education. Analyses of variance were computed separately for both points of measurement.

Prior to these analyses of variance, the homogeneity of variances between the three experimental conditions was tested with the Levene test (Levene, 1960). If the null hypothesis of variance homogeneity was rejected ($p < .05$), Fisher's least significant difference test (LSD; Fisher, 1935) was computed as a post-hoc test to reveal the differences between experimental conditions, as this test is robust with regard to a violation of variance homogeneity. If the null hypothesis of variance of homogeneity was retained ($p > .05$), the more conservative Scheffé test (Scheffé, 1959) was computed to reveal the differences between experimental conditions. This procedure was used within all following analyses of variance and covariance.

To test if the experimental implementation of a shared goal was successful and if it led to a higher level of goal representation and goal commitment in the shared-goal condition than in the other two conditions, analyses of variance with experimental condition as independent variable and the measures of goal representation and goal commitment (see Section 4.3.1.7) as dependent variables were computed. Further, paired *t* tests were applied to investigate if the intrinsic motivation of the participants within the shared goal condition was higher than their extrinsic motivation, and if their cooperative orientation was higher than their competitive orientation (see Section 4.3.1.8).

To determine control variables for the following analyses, associations between all measures of empathy and control measures as personality traits (see Section 4.3.1.3, 4.3.1.4, and 4.3.1.5) and the main demographic characteristics gender (see Section 2.3.1), age (see Section 2.3.2), and education (Feshbach & Feshbach, 2009) were investigated by bivariate correlations, multiple regressions, and univariate ANOVAs. Considering the finding by O'Brien et al. (2013; see Section 2.3.2), non-linear associations between age and empathy measures were additionally investigated. Control measures were included as covariates or

additional predictors in the following analyses, if there was a significant association between the control measure and an empathy measure at one point of measurement.

In a first step towards answering the first research question, partial correlations (a) between individual and systemic measures, (b) between different individual measures, and (c) between different systemic measures of empathy were computed, following the multitrait-multimethod approach (Campbell & Fiske, 1959). As a Fisher z-transformation (Fisher, 1915), the subsequent aggregation of the numerous correlation coefficients (as proposed by Corey, Dunlap, and Burke, 1998) and their comparison (Steiger, 1980) is not possible within the present investigation due to the heterogeneity of the applied methods and subscales, an alternative approach was chosen to test the definite differences between individual and systemic empathy and their measurability. Based on the specific multitrait-multimethod model by Eid (2000), a factor analysis was conducted in AMOS 22.0, with Individual and Systemic Empathy as two higher-order factors. As it is further important to differentiate between emotional and cognitive empathy (see Section 2.1.2), two additional factor analyses with the same higher-order factors were conducted, the first only for cognitive empathy measures and the second only for emotional empathy measures. Measures being neither distinct cognitive nor emotional empathy measures (e.g., the Public Goods Game, see Section 4.3.3) could not be included in these additional analyses. To determine if the measurability of systemic empathy is dependent on group formation processes (see Section 4.2.2), the three mentioned factor analyses were computed separately for both points of measurement. The score difference between both points of measurement was not included in these analyses, as this dynamic component is not comparable with static measurements of empathy. Further, static measurements are more important than dynamic components for investigating the measurability of traits as individual and systemic empathy and possible method effects, as well as the associations and differences between them, which is the focus of the first research question.

The overall model fit of all six confirmatory factor analyses was evaluated by three criteria, following a recommendation by Hu and Bentler (1998): The χ^2 -value, representing the magnitude of discrepancy between the observed and expected covariance matrices, the Root Mean Square Error of Approximation (RMSEA) with its 90% confidence interval, and the Comparative Fit Index (CFI), measuring the incremental fit of the specified model over a null model with uncorrelated variables and freely estimated means. RMSEA values below .05 and

CFI values above .95 were regarded as indicators of a good fit between observed and expected covariance matrices, whereas RMSEA values between .05 and .10 and CFI values between .95 and .90 were regarded as indicators of acceptable fit between observed and expected covariance matrices. RMSEA values above .10 and CFI values below .90 were regarded as indicators of poor model fit (Aiken, Stein, & Bentler, 1994; Brown & Cudeck, 1993; Hu & Bentler, 1998).

The alpha-level was set at $p < .05$ for all preliminary analyses.

4.4.3 Testing for the Hypotheses

The effect of the experimental condition on individual measures of empathy was tested by analyses of covariance for repeated measures, with experimental condition as the between-subjects factor, the individual empathy measures at the first and the second point of measurement as the within-subjects factor, and the previously determined control variables as covariates. In addition to the previously mentioned tests (Levene's test, Fisher's LSD, Scheffé test), Mauchly's test of sphericity (Mauchly, 1940) was applied prior to each analysis, as homogeneous variances and correlations between the factor levels are an important condition for analyses of covariance for repeated measures (Crowder & Hand, 1990). If sphericity was violated, the Greenhouse-Geisser correction (Greenhouse & Geisser, 1959) was applied to address the increased probability of a false positive error. To account for the nested data structure, the group affiliation of each participant was included as a random factor within the analysis. The alpha-level was set at $p < .05$ for the analyses of covariance for repeated measures.

The effect of the experimental condition on systemic measures of empathy was also analyzed by analyses of covariance for repeated measures, with experimental condition as between-subjects-factor, the difference between two measurements of systemic empathy as within-subjects factor, and determined control measures as covariates. However, as the investigation of systemic variables demands an additional analysis of group-level effects, and not only individual-level effects (see Section 2.4.1 and Section 2.4.2), an additional multilevel design was chosen for the statistical analysis of systemic measures. By this means, it was possible to differentiate between effects on group level ($N=18$) and individual level ($N=98$) within the present study. As the multilevel analysis is based on a decomposition and

comparison of individual-level and group-level variance, this approach is similar to the already presented latent group model by Gonzales and Griffin (2002).

The already mentioned economic restrictions did not allow an investigation of more than 18 groups in the present investigation. However, Maas and Hox (2005) demonstrated that a sample size of 50 groups or less at the second level of a multilevel model may lead to severe statistical restrictions: such small sample sizes lead to biased estimates of the second-level standard errors and to a reduced power, which results in a high probability of a false-negative error. In fact, according to power tables for multilevel experimental designs (Konstantopoulos, 2009), a sample size of $N=18$ at the second level of a multilevel model allows only the detection of large effects with an effect size of Cohen's $d=.70$ or more at the recommended power-level of $1-\beta=.80$ (Cohen, 1988). A medium effect with a Cohen's d of $.50$ would be detected with a power of $1-\beta=.31$, and a small effect with a Cohen's d of $.30$ would be detected with a power of $1-\beta=.14$. Both values are far from the recommended power-level of $1-\beta=.80$. Nonetheless, multilevel analyses which are based on a small sample size may be methodologically more preferable than ignoring the nested data structure. To address the power problem and the high probability of a false-negative error, the alpha-level was set at $p<.10$ for the multilevel analyses.

4.4.4 Follow-Up Analyses

Multilevel analyses on the public-goods game (see Section 4.3.3) revealed a significant group-level effect of experimental condition on the difference in donated money between the two points of measurement (see Section 5.5.3 for details). As this effect may be related to the findings of Stürmer et al. (2006; see Section 2.4.3), multiple regressions were administered as follow-up analyses to test if this effect is moderated or mediated by control variables as group climate and group identification. Furthermore, multiple regressions were applied to test for a moderating or mediating effect of perceived diversity on the interrelationship between experimental condition and longitudinal changes in empathy.

Regressions were inspected for heteroscedasticity by scatterplots, for multicollinearity by bivariate correlations between the predictors, and for the normal distribution of z-standardized residuals by the Kolmogorov-Smirnov test for normality (Massey, 1951). A violation of the mentioned conditions is subsequently reported in Section 5.6.

5. Results

The first part of the results section gives a general overview over all measured variables, in terms of descriptive statistics, reliabilities, and a preliminary view on differences between the three experimental conditions. The second part provides several analyses to determine if the experimental manipulations in the present investigation were successful. The third part deals with the identification of relevant control variables for the following analyses.

The following parts are presented in order of the research questions (see Section 3). In the fourth part, confirmatory factor analyses with the two higher-order factors individual and systemic empathy were computed to determine the difference between empathy as an individual and as a systemic trait. In the fifth part, the second and third research question, as well as both hypotheses on the impact of diversity and a shared goal on empathy are tested. The sixth and final part provides some additional follow-up analyses on possible moderating, mediating, and suppression effects of systemic traits (e.g., group identification).

5.1 Descriptive Statistics and Reliabilities

5.1.1 Empathy Measures

Table 4 provides the means and standard deviations of all empathy measures for the first experimental session, for the total sample, as well as separately for the three experimental conditions. Univariate ANOVAs were administered to identify *static* significant differences between the three experimental conditions at one point of measurement.

As only seven acts within group discussions were encoded as displays of emotional empathy (compare Section 4.3.4), the DCS scores measuring cognitive and emotional empathy of a sender were merged into one score ("Empathic Statements").

Table 5 depicts the difference (Δ) between means and standard deviations of all empathy measures in the first and the fourth experimental session. Univariate ANOVAs were administered to investigate if the experimental groups significantly differed in the Δ between empathy measures in the first and fourth experimental session.

Table 4

Descriptive Statistics of Empathy Measures for the First Experimental Session

Measures	Mean (Standard Deviation)				<i>p</i>
	Total Sample <i>N</i> = 98	Control Condition <i>n</i> = 34	Low Diversity <i>n</i> = 31	Shared Goal <i>n</i> = 33	Difference Between Conditions
<i>IRI</i> ^a					
Cognitive Empathy	3.44 (.48)	3.40 (.45)	3.46 (.50)	3.47 (.49)	.79
- Perspective Taking	3.54 (.68)	3.49 (.73)	3.54 (.68)	3.61 (.63)	.77
- Fantasy	3.31 (.66)	3.24 (.63)	3.36 (.69)	3.34 (.69)	.75
Emotional Empathy	3.13 (.42)	3.18 (.46)	3.14 (.47)	3.07 (.35)	.59
- Empathic Concern	3.48 (.51)	3.48 (.53)	3.39 (.53)	3.58 (.47)	.32
- Personal Distress	2.77 (.63)	2.86 (.61)	2.89 (.68)	2.56 (.57)	.07 ⁺
<i>CEEQ-GROUP</i> ^a					
Cognitive Empathy	3.23 (.53)	3.19 (.58)	3.23 (.44)	3.27 (.56)	.85
- Perspective Taking	3.29 (.57)	3.24 (.57)	3.28 (.58)	3.34 (.58)	.75
- Mental Perception	3.19 (.64)	3.16 (.71)	3.21 (.51)	3.20 (.69)	.95
Emotional Empathy	3.20 (.51)	3.23 (.50)	3.10 (.53)	3.25 (.51)	.48
- Empathic Concern	3.52 (.60)	3.55 (.61)	3.40 (.57)	3.60 (.61)	.39
- Mirroring	2.92 (.55)	2.96 (.55)	2.85 (.60)	2.94 (.51)	.69
<i>MET</i>					
Cognitive Empathy ^b	.50 (.17)	.50 (.16)	.48 (.18)	.52 (.17)	.64
Emotional Empathy ^c	4.92 (1.20)	4.98 (1.21)	4.75 (1.01)	5.03 (1.36)	.63
<i>Public Goods Game</i>					
Given Amount in € ^d	2.83 (1.21)	2.85 (1.26)	2.76 (1.12)	2.90 (1.27)	.90
<i>DCS</i>					
Emp. Statements ^e	.17 (.20)	.19 (.19)	.14 (.20)	.17 (.21)	.65
Empathic Feedback ^e	.41 (.33)	.40 (.34)	.41 (.34)	.41 (.30)	.99

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed).

^a Scale ranging from 1(="totally disagree") to 5(="totally agree")

^b Scale ranging from 0(="false") to 1(="true")

^c Scale ranging from 1(="not at all") to 9 (="very much")

^d Scale ranging from 0 to 4 €

^e Scale ranging from 0(="no empathy") to 3(="explicit empathy")

Table 5

Difference (Δ) Between Means and Standard Deviations in the First and Fourth Session

Measures	Δ Between Means (Standard Deviations)				<i>p</i>
	Total Sample <i>N</i> = 91	Control Condition <i>n</i> = 30	Low Diversity <i>n</i> = 31	Shared Goal <i>n</i> = 30	Difference Between Conditions
<i>IRI</i> ^a					
Cognitive Empathy	-.03 (+.05)	-.01 (+.03)	-.08 (+.09)	+.01 (+.05)	.62
- Perspective Taking	-.01 (-.05)	+.03 (-.16)	-.04 (+.01)	-.01 (+.00)	.83
- Fantasy	-.03 (+.00)	+.02 (-.08)	-.12 (+.08)	+.03 (-.01)	.49
Emotional Empathy	-.07 (+.03)	-.03 (+.02)	-.09 (+.00)	-.10 (+.05)	.62
- Empathic Concern	-.06 (+.07)	-.03 (+.10)	-.08 (+.08)	-.07 (+.01)	.87
- Personal Distress	-.10 (-.01)	-.03 (-.07)	-.14 (+.01)	-.13 (-.02)	.60
<i>CEEQ-GROUP</i> ^a					
Cognitive Empathy	+.04 (-.01)	+.01 (+.04)	+.03 (+.07)	+.07 (-.04)	.86
- Perspective Taking	+.05 (-.06)	+.12 (-.04)	+.07 (-.05)	-.03 (-.11)	.46
- Mental Perception	+.01 (+.04)	-.08 (+.05)	-.01 (+.09)	+.14 (+.00)	.23
Emotional Empathy	+.00 (+.01)	+.05 (+.03)	-.05 (-.01)	+.01 (-.01)	.61
- Empathic Concern	-.01 (+.01)	-.05 (-.01)	-.06 (+.00)	+.08 (+.01)	.47
- Mirroring	+.01 (+.00)	+.13 (+.00)	-.04 (-.02)	-.05 (+.04)	.26
<i>MET</i>					
Cognitive Empathy ^b	+.02 (+.03)	+.01 (+.04)	+.02 (+.01)	+.03 (+.03)	.66
Emotional Empathy ^c	+.09 (+.29)	-.06 (+.13)	+.08 (+.12)	+.24 (+.18)	.46
<i>Public Goods Game</i>					
Given Amount in € ^d	+.07 (+.25)	-.38 (+.36)	+.05 (+.29)	+.53 (-.09)	.02 *
<i>DCS</i>					
Emp. Statements ^e	-.02 (-.01)	-.07 (-.06)	+.05 (+.01)	-.04 (-.01)	.17
Empathic Feedback ^e	+.01 (+.04)	+.10 (+.01)	-.05 (+.06)	-.03 (+.04)	.21

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed).

^a Scale ranging from 1(="totally disagree") to 5(="totally agree")

^b Scale ranging from 0(="false") to 1(="true")

^c Scale ranging from 1(="not at all") to 9 (="very much")

^d Scale ranging from 0 to 4 €

^e Scale ranging from 0(="no empathy") to 3(="explicit empathy")

As given in Table 4, means and standard deviations of the DCS were extremely low considering the range of the used scale. Further, Table 5 depicts a low Δ between the DCS-values of the first and fourth experimental session. Both results may indicate that explicit empathic communication and feedback did rarely occur among group members within all analyzed group discussions.

Univariate ANOVAs did not identify any initial significant differences between the three experimental conditions on empathy measures. Such an initial significant difference could not have been caused by the intervention program in the second and third experimental session and would state a serious problem for the main analyses, as such a difference may indicate that the sample of the present investigation is not a random sample but initially biased with regard to empathy. There was only one statistical trend towards a significant difference between the experimental conditions on the IRI Subscale Personal Distress ($F(1,97)=2.81$; $p=.065$; $\eta^2=.06$). However, after including Social Desirability as a covariate (Social Desirability differed significantly between the three experimental conditions, see Section 5.1.2), the statistical trend disappeared ($F(1,90)=1.31$; $p=.31$; $\eta^2=.09$).

The difference in scores between the first and the fourth experimental session was in almost all cases moderate and did not significantly differ between the three experimental conditions, with one exception: The difference between the amount of money that was paid into the common cash box within the Public Goods Game in the first and in the fourth experimental session differed significantly among the experimental groups ($F(1,90)=4.37$; $p=.015$; $\eta^2=.07$). A Scheffé post-hoc test revealed a significant difference between the control and the shared goal condition ($p=.015$), but no difference between the control and the low diversity condition ($p=.34$) or between the low diversity and the shared goal condition ($p=.33$). This result is a first indicator for a hypothesis-confirming effect and will be further evaluated within the main analyses.

To determine the reliability of empathy measures, Cronbach's Alpha (Cronbach, 1951) was computed as an indicator of internal consistency for all empathy questionnaires and the MET, and bivariate correlations between the first and the second point of measurement were computed as indicators of re-test reliability for every empathy measure. The results of the reliability analysis can be derived from Table 6.

Table 6

Reliability of Empathy Measures

Measures	<i>Internal Consistency: First Session (N=98)</i>	<i>Internal Consistency: Fourth Session (N=91)</i>	<i>Re-Test Reliability: Three Weeks</i>
<i>IRI</i>			
Cognitive Empathy	.73	.78	.65
- Perspective Taking	.81	.74	.73
- Fantasy	.66	.77	.38
Emotional Empathy	.67	.69	.65
- Empathic Concern	.63	.66	.75
- Personal Distress	.72	.71	.57
<i>CEEQ-GROUP</i>			
Cognitive Empathy	.83	.85	.71
- Perspective Taking	.71	.66	.67
- Mental Perception	.82	.87	.69
Emotional Empathy	.79	.80	.72
- Empathic Concern	.72	.71	.70
- Mirroring	.63	.66	.61
<i>MET</i>			
Cognitive Empathy	.75	.78	.84
Emotional Empathy	.91	.94	.79
<i>Public Goods Game</i>			
Given Amount in €	-	-	.47
<i>DCS</i>			
Emp. Statements	-	-	.18
Empathic Feedback	-	-	.53

As given in Table 6, the internal consistencies of all empathy measures were at least acceptably high. Several scales of the systemic empathy measure CEEQ-GROUP proved to be highly reliable with values of $\alpha = .80$ or higher. However, despite the short re-test interval of only three weeks, the stability of some empathy measures was low, especially the stability of the IRI subscales Fantasy ($r = .38$) and Personal Distress ($r = .57$), as well as the stability of the

Public Goods Game ($r=.47$) and parameters of empathic communication as measured with the DCS ($r=.18$ for Empathic Statements, $r=.53$ for Empathic Feedback). This may imply that these measures reflect rather empathic states than empathic traits, or that these measures are highly dependent on general emotional states (see Section 6.1.3 for a detailed discussion). Nonetheless, these measures were not excluded within the following analyses, as empathic states may be also of interest with regard to the research questions of the present investigation.

5.1.2 Control Measures

SDS

Based on a true (=1) / false (=0) response format, the Social Desirability Scale (Stöber, 1999; 2001) had a mean of .55 ($SD=.19$; Range: .06-.94) in the total sample ($N=91$). Within the control condition ($N=30$), the mean was .56 ($SD=.21$; Range: .17-.94), within the low diversity condition ($N=31$), the mean was .46 ($SD=.13$; Range: .06-.72), and within the shared goal condition ($N=30$), the mean was .62 ($SD=.18$; Range: .11-.89).

A univariate analysis of variance revealed that SDS-scores differed significantly between the experimental conditions ($F(1,90)=6.30$; $p=.003$; $\eta^2=.11$). Therefore the SDS was included within all main analyses as a covariate, including the multilevel analyses (Section 5.5). A Scheffé post-hoc test revealed a significant difference between the low diversity and the shared goal condition ($p=.003$), but no difference between the control and the low diversity condition ($p=.12$), or between the control and the shared goal condition ($p=.36$).

The internal consistency of the SDS was acceptable with $\alpha=.70$. This value is comparable to the internal consistency reported by Stöber (2001; $\alpha=.72$).

NEO-FFI

Means and standards deviations of the NEO-FFI are depicted in Table 7. Univariate analyses of variance did not reveal any significant differences between the three experimental conditions on the five NEO-FFI subscales.

The internal consistency of all five subscales of the NEO-FFI was satisfactorily high (Neuroticism: $\alpha=.84$; Extraversion: $\alpha=.87$; Openness: $\alpha=.78$; Conscientiousness: $\alpha=.86$; Agreeableness: $\alpha=.74$). These internal consistencies are comparable to the reliabilities reported by Borkenau and Ostendorf (1993).

Table 7

Means and Standard Deviations of the NEO-FFI

Scale	Mean (Standard Deviation)				<i>p</i>
	Total Sample <i>N</i> = 91	Control Condition <i>n</i> = 30	Low Diversity <i>n</i> = 31	Shared Goal <i>n</i> = 30	
Neuroticism	2.89 (.65)	2.97 (.56)	2.94 (.69)	2.75 (.69)	.36
Extraversion	3.34 (.67)	3.35 (.61)	3.47 (.71)	3.20 (.69)	.30
Openness	3.74 (.60)	3.68 (.54)	3.72 (.62)	3.82 (.63)	.65
Conscientiousness	3.70 (.50)	3.74 (.50)	3.56 (.54)	3.80 (.45)	.16
Agreeableness	3.60 (.63)	3.59 (.59)	3.50 (.67)	3.71 (.62)	.46

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed).

The response format ranged from 1="totally disagree" to 5="totally agree".

Questionnaire on Group Processes

The means and standard deviations of the Questionnaire on Group Processes measuring group identification and group climate can be derived from Table 8. The Δ between the means and standard deviations of the subscale Group Identification in the first and the fourth experimental sessions was +.13 (+.08) within the total sample, -.08 (+.39) within the control condition, +.23 (-.12) within the low diversity condition, and +.24 (-.05) within the shared goal condition. However, the difference between the Δ of the three experimental conditions was not significant ($F(1,90)=.71$; $p=.50$; $\eta^2=.02$).

A univariate ANOVA revealed a statistical trend towards a global significant difference between the three experimental conditions in group climate in the second session ($F(1,93)=2.47$; $p=.09$; $\eta^2=.05$). However, the statistical trend disappeared after including Social Desirability as a covariate ($F(1,93)=.97$; $p=.38$; $\eta^2=.02$). A further univariate ANOVA detected a significant difference between the experimental conditions in group climate in the fourth session ($F(1,90)=5.34$; $p=.006$; $\eta^2=.09$). This effect was persistent even after including Social Desirability as a covariate ($F(1,90)=5.24$; $p=.007$; $\eta^2=.08$). A Scheffé post-hoc test revealed that persons in the shared goal condition scored significantly higher on Group Climate than persons in the control condition ($p=.01$). The difference between the low diversity condition and the control condition was marginally significant ($p=.06$). The low diversity condition and the shared goal condition did not differ in Group Climate ($p=.75$).

Table 8

Means and Standard Deviations of the Questionnaire on Group Processes

Scale	Mean (Standard Deviation)				<i>p</i>
	Total Sample	Control Condition	Low Diversity	Shared Goal	Difference Between Conditions
<i>First Session</i> ^a					
Group Identification	2.92 (.83)	3.03 (.69)	2.75 (.93)	2.97 (.87)	.24
Group Climate	3.88 (.62)	3.94 (.60)	3.87 (.65)	3.82 (.63)	.71
<i>Second Session</i> ^b					
Group Identification	3.34 (.70)	3.34 (.61)	3.27 (.76)	3.40 (.74)	.72
Group Climate	4.23 (.66)	4.37 (.57)	4.03 (.78)	4.30 (.59)	.09 ⁺
<i>Third Session</i> ^c					
Group Identification	3.25 (.80)	3.32 (.85)	3.32 (.81)	3.11 (.74)	.66
Group Climate	4.26 (.76)	4.18 (.96)	4.33 (.57)	4.28 (.72)	.56
<i>Fourth Session</i> ^c					
Group Identification	3.05 (.91)	2.95 (1.08)	2.98 (.81)	3.21 (.82)	.39
Group Climate	4.02 (.92)	3.63 (1.32)	4.10 (.65)	4.31 (.49)	.006 ^{**}

Note. ⁺ $p < .10$; ^{*} $p < .05$; ^{**} $p < .01$; ^{***} $p < .001$ (two-tailed). ^a $N=99$, ^b $N=94$, ^c $N=91$. The response format ranged from 1="totally disagree" to 5="totally agree".

Further, the Δ between the means and standard deviations of the subscale Group Climate in the first and the fourth experimental session (total sample: $\Delta M = +.14$, $\Delta SD = +.30$; control condition: $\Delta M = -.31$, $\Delta SD = +.72$; low diversity condition: $\Delta M = +.23$, $\Delta SD = +.00$; shared goal condition: $\Delta M = +.49$, $\Delta SD = -.14$) differed significantly between the three experimental conditions ($F(1,90) = 5.25$; $p = .007$; $\eta^2 = .09$).

Both results indicate that the experimental procedure elicited recognizable changes of group climate in the low diversity condition and in the shared goal condition, but not in the control condition. As a consequence, group climate was included as a covariate within the main analyses. Further, a mediating effect of group climate on the interrelationship between group parameters and empathy measures is investigated in Section 5.6.

The internal consistencies of both subscales were sufficiently high (Group Identification: $\alpha=.77$ in the first session, $\alpha=.75$ in the second session, $\alpha=.83$ in the third session, $\alpha=.88$ in the fourth session; Group Climate: $\alpha=.84$ in the first session, $\alpha=.86$ in the second session, $\alpha=.91$ in the third session, and $\alpha=.95$ in the fourth session). The three-week re-test reliability of both subscales was about zero ($r=.00$ for Group Identification, $r=.04$ for Group Climate), indicating both constructs to be pure states with no systemic trait qualities at all. This result is in accordance with previous research findings (see Section 4.3.1.5).

Group Composition, Goal Representation, and Task Motivation and Orientation

As the questionnaires on Group Composition (Section 4.3.1.6), Goal Representation (Section 4.3.1.7), and Task Motivation and Orientation (Section 4.3.1.8) were administered to control the effectivity of the experimental manipulation, descriptive statistics and reliabilities of these measures are reported together with concurrent univariate ANOVAS in Section 5.2 (“*Manipulation Check*”).

DCS: Dominance and Affiliation

In the first experimental session, the subscale Dominance (measured with the DCS; see Section 4.3.4) had a mean of 3.16 ($SD=.25$) within the total sample ($N=98$), a mean of 3.14 ($SD=.20$) within the control condition ($N=34$), a mean of 3.18 ($SD=.31$) within the low diversity condition ($N=31$), and a mean of 3.16 ($SD=.22$) within the shared goal condition ($N=33$). These scores did not significantly differ between the three experimental conditions ($F(1,97)=.25$; $p=.78$; $\eta^2=.01$). The subscale Affiliation had a mean of 3.20 ($SD=.22$) within the total sample ($N=98$), a mean of 3.22 ($SD=.19$) within the control condition ($N=34$), a mean of 3.17 ($SD=.24$) within the low diversity condition ($N=31$), and a mean of 3.22 ($SD=.22$) within the shared goal condition ($N=33$). There was also no significant difference between the experimental conditions on the subscale Affiliation in the first experimental session ($F(1,97)=.59$; $p=.56$; $\eta^2=.01$).

The Δ between the means and standard deviations of the subscale Dominance in the first and the fourth experimental session was moderate in all experimental conditions (total sample: $\Delta M=-.03$, $\Delta SD=+.15$; control condition: $\Delta M=-.03$, $\Delta SD=+.13$; low diversity condition: $\Delta M=-.09$, $\Delta SD=+.08$; shared goal condition: $\Delta M=+.00$, $\Delta SD=+.08$) and did not differ significantly between the three experimental conditions ($F(1,90)=.15$; $p=.86$; $\eta^2=.00$). There

was also no significant difference with regard to the Δ between the means and standard deviations of the subscale Affiliation in the first and the fourth experimental session (total sample: $\Delta M = +.01$, $\Delta SD = -.01$; control condition: $\Delta M = -.05$, $\Delta SD = +.02$; low diversity condition: $\Delta M = +.04$, $\Delta SD = -.02$; shared goal condition: $\Delta M = +.04$, $\Delta SD = -.02$; $F(1,90) = .74$; $p = .48$; $\eta^2 = .02$).

The three-week re-test reliability of Dominance was low ($r = .55$), and the three-week re-test reliability of Affiliation was zero ($r = .00$). This indicates that both constructs and especially affiliation within group discussions are states and highly changeable.

5.2 Manipulation Check

5.2.1 Perceived Surface-Level Diversity and Diversity Beliefs

In the present study, a successful experimental manipulation of the group parameter Diversity requires not only a modification of objective indicators as age, gender, and education (which was already confirmed by Levene's test for equality of variances in Section 4.1.2), but especially a modification of perceived surface-level diversity (see Section 2.4.4.1). Perceived surface-level diversity was measured by the questionnaire on Group Composition (see Section 4.3.1.6).

In the first experimental session, the mean of perceived surface-level diversity was $M = 2.88$ ($SD = .79$) within the total sample ($N = 98$), $M = 3.19$ ($SD = .71$) within the control condition ($N = 34$), $M = 2.34$ ($SD = .77$) within the low diversity condition ($N = 31$), and $M = 3.06$ ($SD = .63$) within the shared goal condition ($N = 33$). This difference between the three experimental conditions was highly significant ($F(1,97) = 15.71$; $p < .001$; $\eta^2 = .23$). Specifically, a Scheffé post-hoc test revealed that persons within the low diversity condition perceived a lower level of surface-level diversity than persons within the control condition ($p < .001$) and persons within the shared goal condition ($p < .001$). There was no significant difference between the control condition and the shared goal condition ($p = .78$).

In the fourth experimental session, the mean of perceived surface-level diversity was $M = 2.80$ ($SD = .74$) within the total sample ($N = 91$), $M = 3.20$ ($SD = .67$) within the control condition ($N = 30$), $M = 2.22$ ($SD = .56$) within the low diversity condition ($N = 31$), and $M = 3.02$ ($SD = .59$) within the shared goal condition ($N = 30$). The difference between the experimental conditions was also highly significant ($F(1,90) = 26.02$; $p < .001$; $\eta^2 = .36$). Again, a Scheffé post-

hoc test revealed that persons within the low diversity condition perceived a lower level of surface-level diversity than persons within the control condition ($p < .001$) and persons within the shared goal condition ($p < .001$). Again, there was no significant difference between the control condition and the shared goal condition ($p = .58$).

The Δ between the first and the second point of measurement was moderate within all experimental conditions (total sample: $\Delta M = -.08$, $\Delta SD = -.05$; control condition: $\Delta M = +.01$, $\Delta SD = -.04$; low diversity condition: $\Delta M = -.12$, $\Delta SD = -.19$; shared goal condition: $\Delta M = -.04$, $\Delta SD = -.04$) and did not differ between the three experimental conditions ($F(1,90) = 1.08$; $p = .34$; $\eta^2 = .00$). The changes in perceived surface-level diversity between the first and the second point of measurement were investigated by a univariate ANOVA for repeated measures and were neither significant in the total sample ($F(1) = .02$; $p = .89$), nor in any of the three experimental conditions ($F(1) = .38$; $p = .54$ within the control condition, $F(1) = .95$; $p = .33$ within the low diversity condition, and $F(1) = .18$; $p = .67$ within the shared goal condition).

These results are strong indicators of an accurate and successful experimental manipulation, as persons in the low diversity condition perceived lower levels of surface-level diversity within their groups than persons in the other two conditions. This perceptual difference was stable from the very beginning of the experiment until the last point of measurement (the three-week re-test reliability of perceived diversity was $r = .72$).

However, these results are limited by a low internal consistency of the used measure: The subscale Perceived Diversity had an internal consistency of $\alpha = .51$ in the first experimental session and an internal consistency of $\alpha = .49$ in the fourth experimental session. The low internal consistency indicates that participants perceived differences between the general diversity, the diversity with respect to age, the diversity with respect to gender, and the diversity with respect to the ethnic background within their groups. As a consequence, every item of the subscale Perceived Diversity was analyzed separately by a univariate ANOVA. The results of this analysis can be derived from Table 9.

This additional analysis demonstrates that the experimental procedure did not elicit different perceptions of general diversity among the experimental conditions, but rather different perceptions of diversity in the three experimentally manipulated surface-level traits age, gender, and ethnic background. This effect indicates that the item measuring the perception of general diversity was not specific enough and therefore dispensable. However,

the effect is completely in line with the aims of the experimental design. Only 3 out of 18 post-hoc contrasts were not in line with this trend: One insignificant difference between the low diversity and the shared goal condition in the first experimental session with respect to the perceived surface-level diversity in the variable age ($p=.32$), and two statistical trends (low diversity vs. control condition in the first session with respect to the variable age, $p=.07$, and low diversity vs. shared goal condition in the fourth session with respect to the variable ethnic background, $p=.08$). However, both mentioned statistical trends are associated with lower values in perceived surface-level diversity within the low diversity condition (perceived diversity in age within the low diversity condition in the first session: $M=2.11$, $SD=.89$, vs. perceived diversity in age within the control condition in the first session: $M=2.87$, $SD=1.31$; perceived diversity in ethnic background in the low diversity condition in the fourth session: $M=2.13$, $SD=.94$, vs. perceived diversity in ethnic background in the low diversity condition in the fourth session: $M=2.43$, $SD=1.22$).

Table 9

Separate Manipulation Check for all Facets of Perceived Surface-Level Diversity

	<i>Differences Between Conditions</i>				<i>Scheffé Post-Hoc Test</i>		
	<i>F</i>	<i>df</i>	<i>p</i>	η^2	C-LD	C-SG	LD-SG
<i>First Session</i>							
General Diversity	1.31	97	.28	.00	n.s.	n.s.	n.s.
Age	3.10	97	.05	.05	+	n.s.	n.s.
Gender	19.89	97	.000	.30	***	n.s.	***
Ethnic Background	10.08	97	.000	.17	***	n.s.	***
<i>Fourth Session</i>							
General Diversity	1.13	90	.33	.00	n.s.	n.s.	n.s.
Age	5.81	90	.004	.10	*	n.s.	*
Gender	56.37	90	.000	.55	***	n.s.	***
Ethnic Background	5.73	90	.005	.10	**	n.s.	+

Note. + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed).

"C"=Control Condition, "LD"=Low Diversity Condition, "SG"=Shared Goal Condition

The questionnaire's additional subscale Diversity Beliefs had a mean of $M=3.21$ ($SD=.65$) within the total sample ($N=98$), $M=3.24$ ($SD=.84$) within the control condition ($N=34$), $M=3.23$ ($SD=.53$) within the low diversity condition ($N=31$), and $M=3.15$ ($SD=.54$) within the shared goal condition ($N=33$). The three experimental conditions did not significantly differ in diversity beliefs in the first experimental session ($F(1,97)=.13$; $p=.88$; $\eta^2=.00$). The Δ between the means and standard deviations of diversity beliefs in the first and the fourth experimental session was moderate in all experimental conditions (total sample: $\Delta M=-.05$, $\Delta SD=-.07$; control condition: $\Delta M=+.05$, $\Delta SD=-.27$; low diversity condition: $\Delta M=-.15$, $\Delta SD=+.04$; shared goal condition: $\Delta M=-.06$, $\Delta SD=+.06$) and did not differ significantly between the three experimental conditions ($F(1,90)=1.07$; $p=.35$; $\eta^2=.02$).

Finding a significant difference between the three experimental conditions in diversity beliefs would have been for two reasons highly problematic for a successful experimental manipulation in the present investigation: First, such a difference would indicate that the sample may be initially biased with regard to diversity beliefs and therefore cannot represent a random sample. Second, as diversity beliefs moderate various effects of perceived surface-level diversity (see Section 4.1.3.6), a difference between the experimental conditions in diversity beliefs may lead to differently moderated and not controllable effects of perceived surface-level diversity. Therefore, the result that the three experimental conditions did not significantly differ in diversity beliefs is also in line with the aims of the experimental design.

The subscale Diversity Beliefs had a rather low internal consistency (first session: $\alpha=.68$; fourth session: $\alpha=.49$), which indicates the measured construct to be rather heterogeneous. The three-week re-test reliability was moderate with $r=.57$. Out of these reasons, even if diversity beliefs may have a strong impact on various effects of perceived surface-level diversity, no further analyses on moderating effects of the subscale Diversity Beliefs were conducted.

In summary, all analyses that were presented in this section speak for an accurate and successful experimental manipulation of perceived surface-level diversity in the three characteristics age, gender, and ethnic background.

5.2.2 Goal Representation and Goal Commitment

An experimental implementation of a shared goal within a small group can only be regarded as successful and effective, if the specific shared goal is cognitively represented within the corresponding experimental condition and if the participants in this condition are committed to this goal. These constructs have been measured within the questionnaire on Goal Representation (see Section 4.3.1.7).

In the first experimental session, the first item measuring a general goal representation had a mean of $M=3.18$ ($SD=1.42$) within the total sample ($N=98$), $M=2.91$ ($SD=1.36$) within the control condition ($N=34$), $M=2.39$ ($SD=1.33$) within the low diversity condition ($N=31$), and $M=4.21$ ($SD=.86$) within the shared goal condition ($N=33$). The difference between the three experimental conditions was highly significant ($F(1,97)=19.71$; $p<.001$; $\eta^2=.29$). A subsequent Scheffé post-hoc test revealed that persons within the shared goal condition had significantly higher levels of goal representation than persons in the control condition ($p<.001$) and persons in the low diversity condition ($p<.001$). There was no significant difference between the control condition and the low diversity condition ($p=.22$).

In the fourth experimental session, the item on goal representation had a mean of $M=3.25$ ($SD=1.38$) within the total sample ($N=91$), $M=2.67$ ($SD=1.30$) within the control condition ($N=30$), $M=2.74$ ($SD=1.32$) within the low diversity condition ($N=31$), and $M=4.37$ ($SD=.72$) within the shared goal condition ($N=30$). Again, the difference between the three experimental conditions was highly significant ($F(1,90)=21.18$; $p<.001$; $\eta^2=.31$). Another Scheffé post-hoc test was administered and revealed that persons within the shared goal condition had significantly higher levels of goal representation than persons in the control condition ($p<.001$) and persons in the low diversity condition ($p<.001$). Again, there was no significant difference between the control condition and the low diversity condition ($p=.97$).

Participants were asked to specify the represented goal in an open response format (see Section 4.3.1.7). Within the shared goal condition, 16 of 33 participants (48.5%) specified the collage as the shared goal of their group in the first experimental session, and 21 of 30 participants (70%) specified the collage as the shared goal in the fourth experimental session. This percentage increase between the first and the fourth experimental session was significant ($\chi^2(1,N=30)=12.86$; $p<.001$). In the other two experimental conditions, no participant described the individual collage as a shared goal. The

percentage difference between conditions was also significant ($\chi^2(1, N=98)=42.12$; $p<.001$ in the first session, and $\chi^2(1, N=91)=58.19$; $p<.001$ in the fourth session).

In the first experimental session, out of all 82 participants in the total sample who did not specify the collage as the shared goal, 54 participants did not specify any particular shared goal. The remaining 28 participants specified other shared goals, e.g. gaining 60 euro for the participation in the experiment, which was described by 12 participants (17.6% of the total sample) as the shared goal of their group. In the fourth experimental session, out of all 70 participants in the total sample who did not specify the collage as the shared goal of their group, 41 participants did not specify any particular shared goal. The remaining 29 specified other goals. Eighteen participants (19.78% of the total sample) described gaining money for the participation in the experiment as the primary shared goal of their group.

These results indicate that participants in the shared goal condition had significantly higher levels of goal representation than participants in the other two experimental conditions, and that these levels of goal representation were stable throughout the course of the experiment. Further, increasingly more participants in the shared goal condition shared the representation of the collage as the central goal of the group towards the end of the experiment. So far, these findings speak for a successful experimental implementation of a specific shared goal within the present study. However, as several participants reported that gaining 60 euro for the participation in the experiment is the shared goal of their group, it is possible that gaining money as a shared goal interfered with the experimentally implemented goal of creating a creative collage.

If participants specified a shared goal, they were asked to report their level of commitment to this particular goal. Forty-four participants completed the subscale on goal commitment in the first experimental session, and 50 participants completed this subscale in the fourth experimental session.

In the first experimental session, participants in the total sample ($N=44$) reported a mean goal commitment of $M=3.37$ ($SD=.82$), participants within the control condition ($N=12$) a mean of $M=2.88$ ($SD=.76$), participants within the low diversity condition ($N=9$) a mean of $M=3.48$ ($SD=.74$), and participants within the shared goal condition ($N=23$) a mean of $M=3.58$ ($SD=.79$). The difference between the three experimental conditions was significant

($F(1,43)=3.40$; $p=.043$; $\eta^2=.10$). A Scheffé post-hoc test revealed that persons within the shared goal condition had significantly higher levels of goal commitment than persons in the control condition ($p=.048$). There was no difference in goal commitment between the shared goal and the low diversity condition ($p=.95$). There was also no significant difference between the low diversity and the control condition ($p=.22$).

In the fourth experimental session, participants in the total sample ($N=50$) reported a mean goal commitment of $M=3.56$ ($SD=.68$), participants within the control condition ($N=14$) a mean of $M=3.23$ ($SD=.63$), participants within the low diversity condition ($N=13$) a mean of $M=3.86$ ($SD=.69$), and participants within the shared goal condition ($N=23$) a mean of $M=3.59$ ($SD=.65$). These scores differ significantly from each other ($F(1,49)=3.21$; $p=.049$; $\eta^2=.08$), but a Scheffé post-hoc test revealed only one marginally significant difference between the low diversity and the control condition ($p=.052$).

The internal consistency of the subscale Goal Commitment was high ($\alpha=.90$ in the first and in the fourth experimental session) and its three-week re-test reliability near zero ($r=.10$). The low stability of Goal Commitment may reflect the reported significant changes in participants' perceptions of the central shared goal of their group between the first and the fourth experimental session.

Obviously, if participants perceived a shared goal within their group, the levels of goal commitment did not greatly differ between the three experimental conditions. This is an argument for the already mentioned interference of other important shared goals as gaining money with the experimentally implemented goal. However, only few persons in the control condition (12 out of 34 in the first session, 14 out of 30 in the fourth session) and in the low diversity condition (9 out of 31 in the first session, 13 out of 31 in the fourth session) perceived a shared goal at all, in contrast to the shared goal condition (23 out of 33 in the first session, 23 out of 30 in the fourth session). This distribution differed significantly between the experimental conditions at both points of measurement ($\chi^2(1, N=98)=12.63$; $p=.002$ in the first session, $\chi^2(1, N=91)=6.95$; $p=.031$ in the fourth session) and is also reflected by the significantly higher levels of goal representation in the shared goal condition than in the other two experimental conditions. Therefore, the experimental implementation of the shared goal in the present study can be regarded as generally successful.

5.2.3 Task Motivation and Orientation

The experimental implementation of a shared goal in the present study can be regarded as successful and ecologically valid, if participants are intrinsically rather than extrinsically motivated to pursue this goal and if they engage in cooperative rather than competitive behavior (see Sections 2.4.4.2 and 4.2). Intrinsic and extrinsic motivation as well as cooperative and competitive orientation were measured in the shared goal condition with the questionnaire on Task Motivation and Orientation after the second, third, and fourth experimental session (see Section 4.3.1.8). The items of this questionnaire were explicitly related to the work on the creative collage. Out of this reason, interfering effects of other perceived shared goals which were described in the previous section (e.g., gaining money) can be excluded.

The means and standard deviations of this questionnaire are presented in Table 10. A paired *t* test was administered to investigate if participants' intrinsic motivation was significantly higher than their extrinsic motivation, and if cooperative orientation was significantly higher than competitive orientation.

Table 10

Means and Standard Deviations of the Questionnaire on Task Motivation and Orientation

Session	Mean (Standard Deviation)		
	Intrinsic Motivation	Extrinsic Motivation	Difference Between Subscales (paired <i>t</i> test)
<i>Second Session</i> ^a	3.80 (1.05)	1.97 (.74)	<i>t</i> =6.96, <i>p</i> <.001 ***
<i>Third Session</i> ^b	3.73 (.91)	2.29 (.80)	<i>t</i> =5.54, <i>p</i> <.001 ***
<i>Fourth Session</i> ^c	3.78 (.88)	2.18 (.80)	<i>t</i> =6.70, <i>p</i> <.001 ***
	Cooperative Orientation	Competitive Orientation	Difference Between Subscales (paired <i>t</i> test)
<i>Second Session</i> ^a	4.02 (.92)	1.33 (.63)	<i>t</i> =13.61, <i>p</i> <.001 ***
<i>Third Session</i> ^b	4.24 (.75)	1.33 (.56)	<i>t</i> =16.24, <i>p</i> <.001 ***
<i>Fourth Session</i> ^c	4.22 (.81)	1.31 (.55)	<i>t</i> =15.92, <i>p</i> <.001 ***

Note. ⁺ *p* < .10; * *p* < .05; ** *p* < .01; *** *p* < .001 (two-tailed). ^a *N*=33, ^b *N*=31, ^c *N*=30.

The results indicate that participants in the shared goal condition had significantly higher levels of intrinsic than extrinsic motivation and significantly higher levels of cooperative than competitive task orientation. This is another argument for a successful and ecologically valid experimental implementation of a shared goal within the small groups of this experimental condition.

The internal consistencies and two-week re-test reliabilities of this questionnaire can be derived from Table 11. The internal consistencies of the subscale Extrinsic Motivation were low, but still acceptable. The measured re-test reliabilities indicate a sufficient stability in task motivation and orientation throughout the investigation period.

Table 11

Reliability of the Questionnaire on Task Motivation and Orientation

Scale	Internal Consistency			
	Second Session	Third Session	Fourth Session	Two-Week Re-Test Reliability
Intrinsic Motivation	.87	.85	.90	.75
Extrinsic Motivation	.61	.66	.56	.71
Cooperative Orientation	.91	.89	.92	.62
Competitive Orientation	.84	.86	.86	.77

5.3 Associations Between Empathy Measures and Control Variables

Correlational analyses between empathy measures and control variables comprised nine different comparisons: (1) between empathy measures in the first session and control variables in the first session, (2) between empathy measures in the fourth session and control variables in the fourth session, (3) between ΔM of empathy measures and ΔM of control variables, (4) between empathy measures in the first session and control variables in the fourth session, (5) between empathy measures in the fourth session and control variables in the first session, (6) between empathy measures in the first session and ΔM of control variables, (7) between empathy measures in the fourth session and ΔM of control variables, (8) between ΔM of empathy measures and control variables in the first session, and finally (9) between ΔM of empathy measures and control variables in the fourth session.

Only significant correlations are reported in this Section, complete correlation tables are presented in the Appendix (Section 8.2). Not significant or marginally significant correlations are reported in some cases, if an association between an empathy measure and a control variable was strikingly inconsistent between the first and the fourth experimental session.

As Social Desirability and the Big Five personality traits were only measured once, these variables were only correlated with (1) empathy measures in the first experimental session, (2) empathy measures in the second experimental session, and (3) the Δ between means of empathy measures in the first and fourth experimental session.

The measures on perceived diversity, diversity beliefs, and goal representation (see Section 5.2) were also included as control variables within correlational analyses. The measures on goal commitment, task motivation, and task orientation were not analyzed, as only a minority of participants rated the corresponding items and therefore a consideration of these measures as control variables within the total sample would be not applicable.

Further, interrelationships between empathy and the three main demographical characteristics age, gender, and education are presented in Section 5.3.6. Possible linear associations between age and empathy were investigated by bivariate correlations and non-linear associations by multiple regressions with age and squared age as predictors. Associations between gender and empathy were investigated by univariate ANOVAs with gender as independent variable and all empathy subscales as dependent variables. The interrelationship between empathy and education was investigated by bivariate correlations.

If analyses revealed a significant or marginally significant association between an empathy scale and a control variable, this variable was automatically included as a covariate or as an additional predictor within all following analyses on this empathy scale.

5.3.1 Social Desirability

Social desirability was significantly correlated with the IRI subscale Perspective Taking in the fourth experimental session ($r=.23$, $p=.027$). In the first experimental session, the association between social desirability and the IRI subscale Perspective Taking was marginally significant ($r=.19$, $p=.066$). Further, high levels of social desirability were negatively associated with the IRI subscale Personal Distress in the first ($r=-.23$, $p=.028$) and

in the fourth ($r=-.23, p=.032$) experimental session. There was also a significant association between social desirability and the CEEQ-GROUP subscale Perspective Taking, which is a systemic empathy measure, but only in the fourth experimental session ($r=.21, p=.043$), not in the first experimental session ($r=.14, p=.20$). Moreover, the analyses revealed a significant negative correlation between social desirability and the MET subscale Cognitive Empathy in the first session ($r=-.32, p=.002$). The same interrelationship was marginally significant in the fourth experimental session ($r=-.21, p=.051$). The DCS group discussion parameter Empathic Feedback was positively associated with social desirability in the fourth session ($r=.22, p=.039$), but not in the first session ($r=-.06, p=.61$). With regard to the Δ between means of empathy measures in the first and fourth experimental session, high levels of social desirability were positively associated with an increase of systemic mental state perception (CEEQ-GROUP subscale; $r=.28, p=.006$) and increased levels of the DCS group discussion parameter Empathic Feedback ($r=.27, p=.009$).

5.3.2 Big Five Personality Traits

Interestingly, correlational analyses revealed multiple associations between all Big Five personality traits and various empathy measures.

Neuroticism was substantially and significantly associated with the IRI subscale Personal Distress, both in the first ($r=.36, p<.001$) and in the fourth experimental session ($r=.48, p<.001$). Further, Neuroticism was significantly associated with the IRI subscale Empathic Concern in the fourth session ($r=.26, p=.015$), but not in the first session ($r=.11, p=.29$). Moreover, Neuroticism was negatively associated with the amount of money that was paid into the common cash box within the Public Goods Game in the fourth session ($r=-.30, p=.004$), but there was no such association in the first session ($r=-.14, p=.18$). Neuroticism was also significantly associated with the ΔM of several empathy measures between the first and fourth session: Higher levels of Neuroticism were associated with an increase of individual empathic concern (IRI subscale; $r=.23, p=.028$), but a decrease in systemic empathic concern (CEEQ-GROUP subscale; marginally significant with $r=-.20, p=.053$), and with a decrease in the amount of money that was paid into the common cash box within the Public Goods Game ($r=-.21, p=.050$).

Extraversion was positively associated with the IRI subscale Fantasy in the first session ($r=.21, p=.044$), but not in the fourth session ($r=.03, p=.79$). Analyses revealed a

further significant association between Extraversion and the systemic CEEQ-GROUP subscale Mirroring in the fourth session ($r=.24, p=.022$), but there was no such association in the first session ($r=.05, p=.66$). Extraversion was negatively associated with the MET subscale Cognitive Empathy both in the first ($r=-.28, p=.007$) and in the fourth experimental session ($r=-.25, p=.019$). Finally, Extraversion was associated with a decrease in individual empathic concern between the first and the fourth experimental session (IRI subscale; $r=-.24, p=.025$), and an increase in systemic mirroring (CEEQ-GROUP subscale, $r=.23, p=.030$).

Openness was not significantly correlated with any IRI subscale in the first experimental session, but with three IRI subscales in the fourth session: Perspective Taking ($r=.23, p=.026$), Fantasy ($r=.26, p=.013$), and Empathic Concern ($r=.24, p=.020$). Further, Openness was significantly associated with the MET subscale Cognitive Empathy in the first ($r=.30, p=.004$) and in the fourth session ($r=.26, p=.013$). Analyses revealed another interrelationship between Openness and the MET subscale Emotional Empathy in the first ($r=.22, p=.039$), but not in the fourth experimental session ($r=.17, p=.11$). Moreover, Openness was significantly correlated with the DCS group discussion parameter Empathic Statements, in the first ($r=.22, p=.036$) as well as in the fourth session ($r=.36, p<.001$). With regard to the differences between the first and the fourth experimental session, higher levels of Openness were significantly associated with an increase in individual fantasy (IRI subscale; $r=.21, p=.042$).

Conscientiousness was next to Agreeableness the Big Five personality trait with the most remarkable pattern of correlations with empathy measures. With regard to individual empathy measures, there was only one significant association between Conscientiousness and the IRI subscale Perspective Taking in the first ($r=.30, p=.004$) and in the fourth session ($r=.23, p=.032$). The MET subscale Emotional Empathy was significantly associated with Conscientiousness in the fourth session ($r=.24, p=.025$), but not in the first session ($r=.16, p=.11$). With regard to systemic empathy measures, Conscientiousness was significantly associated with the CEEQ-GROUP subscale Perspective Taking in the first ($r=.21, p=.042$) and in the fourth session ($r=.24, p=.021$). However, several systemic empathy measures were associated with Conscientiousness in the fourth session, at least on a marginally significant level, but there was no such association in the first session: The CEEQ-GROUP subscale Mental State Perception ($r=.05, p=.61$ in the first session, $r=.18, p=.082$ in the fourth session), the CEEQ-GROUP subscale Empathic Concern ($r=.15, p=.16$ in the first session,

$r=.25$, $p=.017$ in the fourth session), the CEEQ-GROUP subscale Mirroring ($r=.12$, $p=.25$ in the first session, $r=.18$, $p=.081$ in the fourth session), and the amount of money that was paid into the common cash box within the Public Goods Game ($r=.12$, $p=.26$ in the first session, $r=.24$, $p=.021$ in the fourth session). Further, higher levels of Conscientiousness were associated with increased levels of the DCS group discussion parameter Empathic Statements ($r=.21$, $p=.045$) between the first and the fourth experimental session.

Agreeableness was substantially interrelated with various individual and systemic empathy measures. These interrelationships were mostly consistent at both points of measurement, in contrast to the associations between Conscientiousness and empathy: The IRI subscale Perspective Taking was significantly correlated with Agreeableness both in the first ($r=.32$, $p=.002$) and in the fourth experimental session ($r=.35$, $p=.001$), as well as the IRI subscale Empathic Concern ($r=.30$, $p=.004$ in the first session, $r=.26$, $p=.013$ in the fourth session), the systemic CEEQ-GROUP subscale Empathic Concern ($r=.22$, $p=.038$ in the first session, $r=.30$, $p=.004$ in the fourth session), and the MET subscale Emotional Empathy ($r=.31$, $p=.003$ in the first session, $r=.26$, $p=.012$ in the fourth session). The CEEQ-GROUP subscale Perspective Taking was inconsistently associated with Agreeableness in the fourth session ($r=.22$, $p=.040$) but not in the first session ($r=.16$, $p=.13$), as well as the DCS group discussion parameter Empathic Feedback ($r=.30$, $p=.005$ in the fourth session, $r=.17$, $p=.10$ in the first session). Further, higher levels of Agreeableness were significantly associated with an increase in systemic mental state perception between both points of measurement (CEEQ-GROUP subscale; $r=.31$, $p=.003$).

5.3.3 Group Climate and Identification

The level of group climate in the first experimental session was significantly associated with the CEEQ-GROUP subscale Perspective Taking in the first ($r=.26$, $p=.010$) and in the fourth session ($r=.21$, $p=.047$). Further, there was a significant negative correlation between group climate in the first session and the MET subscale Cognitive Empathy in the first ($r=-.24$, $p=.019$) and in the fourth session ($r=-.27$, $p=.010$). Group climate in the fourth experimental session was significantly associated with the IRI subscale Perspective Taking in the first session ($r=.26$, $p=.014$), the IRI subscale Empathic Concern in the first session ($r=.28$, $p=.008$), the CEEQ-GROUP subscale Perspective Taking in the first session ($r=.21$, $p=.048$), and the CEEQ-GROUP subscale Mental State Perception in the fourth session ($r=.27$, $p=.011$).

Furthermore, the level of group climate in the fourth experimental session was associated with a decrease of individual personal distress (IRI subscale; $r=-.21$, $p=.047$), and an increase in the amount of money that was given into the common cash box within the Public Goods Game ($r=.24$, $p=.022$). An increase in group climate scores between the first and the fourth experimental session was associated with the IRI subscale Empathic Concern both in the first ($r=.26$, $p=.011$) and in the fourth session ($r=.21$, $p=.046$), as well as with the amount of money that was paid into the common cash box within the Public Goods Game ($r=.22$, $p=.032$).

Interestingly, group identification was consistently associated with several individual measures of cognitive empathy, and several systemic measures of emotional empathy. The level of group identification in the first experimental session was significantly associated with the IRI subscale Perspective Taking, both in the first ($r=.25$, $p=.012$) and in the fourth experimental session ($r=.26$, $p=.014$), and with the IRI subscale Fantasy, also both in the first ($r=.23$, $p=.015$) and in the fourth experimental session ($r=.30$, $p=.004$). Further, group identification in the first session was correlated with the CEEQ-GROUP subscale Empathic Concern in the first ($r=.39$, $p<.001$) and in the fourth session ($r=.32$, $p=.002$), and with the CEEQ-GROUP subscale Mirroring in the first ($r=.28$, $p=.005$) and in the fourth session ($r=.46$, $p<.001$). Further associations were revealed between group identification in the first session and the MET subscale Emotional Empathy both in the first ($r=.30$, $p=.002$) and in the fourth session ($r=.22$, $p=.034$). The level of group identification in the fourth experimental session was similarly associated with the IRI subscale Perspective Taking in the first ($r=.27$, $p=.010$) and in the fourth session ($r=.26$, $p=.002$), the IRI subscale Fantasy in the first ($r=.23$, $p=.027$) and in the fourth session ($r=.22$, $p=.034$), the CEEQ-GROUP subscale Empathic Concern in the first ($r=.26$, $p=.013$) and in the fourth session ($r=.35$, $p=.001$), the CEEQ-GROUP subscale Mirroring in the first ($r=.24$, $p=.025$) and in the fourth session ($r=.44$, $p<.001$), and with the MET subscale Emotional Empathy in the first ($r=.21$, $p=.050$) and in the fourth session ($r=.26$, $p=.012$). Group identification in the fourth session was additionally associated with the IRI subscale Personal Distress in the first session ($r=.25$, $p=.019$), the CEEQ-GROUP subscale Mental State Perception in the fourth session ($r=.33$, $p=.002$), the amount of money that was paid into the common cash box within the Public Goods Game in the fourth session ($r=.26$, $p=.013$), the DCS group discussion parameter Empathic Feedback in the fourth session ($r=.23$, $p=.027$), and an increase on the CEEQ-GROUP subscale Mirroring between the first

and the fourth experimental session ($r=.24$, $p=.023$). Finally, an increase of group identification between the first and the fourth session was significantly associated with a decrease of individual personal distress (IRI subscale; $r=-.23$, $p=.032$), as well as with the amount of money that was paid into the common cash box in the first ($r=.25$, $p=.019$) and in the fourth experimental session ($r=.22$, $p=.034$).

5.3.4 Dominance and Affiliation

The DCS group discussion parameter Dominance in the first session was significantly associated with the MET subscale Cognitive Empathy in the first ($r=.41$, $p<.000$) and in the fourth experimental session ($r=.37$, $p=.000$), with the DCS group parameter Empathic Statements in the fourth session ($r=.46$, $p<.001$), and with an increase of the DCS group parameter Empathic Statements between the first and the fourth session ($r=.24$, $p=.025$). The level of dominance in the fourth session was significantly correlated with the DCS group discussion parameter Empathic Statements in the first ($r=.41$, $p<.001$) and in the fourth session ($r=.22$, $p=.032$). Further, an increase in dominance between the first and the fourth experimental session was significantly associated with higher scores on the IRI subscale Empathic Concern in the first session ($r=.25$, $p=.017$) and higher amounts of money that was paid into the common cash box within the Public Goods Game in the fourth session ($r=.22$, $p=.033$).

Interestingly, analyses revealed several negative associations between the DCS group discussion parameter Affiliation in the first session and various empathy measures: the MET subscale Emotional Empathy in the fourth session ($r=-.24$, $p=.024$), the amount of money that was paid into the common cash box within the Public Goods Game in the first ($r=-.27$, $p=.007$) and in the fourth session ($r=-.25$, $p=.019$), and ΔM in systemic mental state perception between the first and the fourth experimental session (CEEQ-GROUP subscale; $r=-.22$, $p=.034$). Further, affiliation in the first session was positively interrelated with the MET subscale Cognitive Empathy in the first ($r=.24$, $p=.019$) and in the fourth session ($r=.26$, $p=.011$). In contrast to this correlation pattern, affiliation in the fourth experimental session was positively associated with the IRI subscale Fantasy in the first ($r=.24$, $p=.020$) and in the fourth session ($r=.28$, $p=.006$), the IRI subscale Empathic Concern in the first ($r=.22$, $p=.031$) and in the fourth session ($r=.26$, $p=.011$), the CEEQ-GROUP subscale Empathic Concern in the fourth session ($r=.31$, $p=.003$), the CEEQ-GROUP subscale Mirroring in the fourth session

($r=.23, p=.032$), the MET subscale Cognitive Empathy in the first session ($r=.25, p=.018$), the MET subscale Emotional Empathy in the first ($r=.32, p=.002$) and in the fourth session ($r=.31, p=.003$), the DCS group discussion parameter Empathic Statements in the fourth session ($r=.49, p<.001$), and further with an increase in systemic perspective taking (CEEQ-GROUP subscale; $r=.25, p=.016$) and an increase in the DCS group discussion parameter Empathic Statements ($r=.39, p<.001$) between the first and the fourth session. Furthermore, an increase in affiliation between the first and the fourth experimental session was significantly associated with an increase in systemic Perspective Taking ($r=.22, p=.036$), an increase in Empathic Statements ($r=.29, p=.005$), the level of Empathic Statements in the fourth session ($r=.30, p=.004$), the IRI subscale Empathic Concern in the first ($r=.30, p=.003$) and in the fourth session ($r=.23, p=.025$), the MET subscale Emotional Empathy in the first ($r=.35, p=.001$) and in the fourth session ($r=.36, p<.001$), and the amount of money that was paid into the common cash box within the Public Goods Game in the fourth session ($r=.26, p=.015$).

5.3.5 Perceived Diversity, Diversity Beliefs, and Goal Representation

Perceived surface-level diversity in the first session was significantly associated with the IRI subscale Empathic Concern in the first ($r=.23, p=.025$) and in the fourth session ($r=.21, p=.048$), the CEEQ-GROUP subscale Empathic Concern in the first session ($r=.26, p=.009$), an increase in individual personal distress (IRI subscale; $r=.27, p=.009$), and a decrease in the amount of money that was paid into the common cash box within the Public Goods Game between the first and the fourth experimental session ($r=-.24, p=.023$). Perceived surface-level diversity in the fourth session was significantly associated with the CEEQ-GROUP subscale Empathic Concern in the first session ($r=.31, p=.003$), the CEEQ-GROUP subscale Mirroring in the first session ($r=.21, p=.044$), an increase in individual personal distress (IRI subscale; $r=.26, p=.014$) and a decrease in systemic empathic concern (CEEQ-GROUP subscale; $r=-.26, p=.026$). Further, an increase in perceived surface-level diversity between both points of measurement was significantly associated with a decrease in systemic perspective taking ($r=-.22, p=.033$) and a decrease in the DCS group discussion parameter Empathic Statements between the first and the fourth session ($r=-.25, p=.019$).

Diversity beliefs in the first experimental session were significantly associated with an increase in individual empathic concern between the first and the fourth experimental

session (IRI subscale; $r=.27, p=.010$). There was no significant association between diversity beliefs in the fourth session and any empathy measure. However, an increase in diversity beliefs was significantly correlated with a decrease in individual personal distress (IRI subscale, $r=-.34, p=.001$) and an increase in systemic perspective taking (CEEQ-GROUP subscale, $r=.26, p=.012$) between the first and the fourth experimental session, as well as with the IRI subscale Personal Distress in the first session ($r=.23, p=.030$).

Goal representation in the first experimental session was significantly associated with the IRI subscale Empathic Concern in the first session ($r=.20, p=.048$), the CEEQ-GROUP subscale Empathic Concern in the fourth session ($r=.22, p=.033$), the MET subscale Cognitive Empathy in the first session ($r=-.22, p=.030$), the MET subscale Emotional Empathy in the fourth session ($r=.30, p=.003$), and further with a decrease in individual empathic concern (IRI subscale; $r=-.22, p=.039$), an increase in systemic empathic concern (CEEQ-GROUP subscale; $r=.28, p=.008$), and an increase of scores on the MET subscale Emotional Empathy ($r=.38, p<.001$) between the first and the fourth session. Goal representation in the fourth session was significantly correlated with the CEEQ-GROUP subscale Mental State Perception in the fourth session ($r=.21, p=.048$), with the MET subscale Cognitive Empathy in the first session ($r=-.22, p=.037$), the amount of money that was paid into the common cash box within the Public Goods Game in the fourth session ($r=.24, p=.024$), a decrease in systemic perspective taking (CEEQ-GROUP subscale; $r=-.28, p=.007$), an increase in scores on the MET subscale Emotional Empathy ($r=.30, p=.004$), and an increase in the amount of money that was paid into the common cash box within the public goods game ($r=.23, p=.027$) between the first and the fourth session. Further, an increase in goal representation between the first and the fourth session was significantly associated with a decrease in systemic empathic concern (CEEQ-GROUP subscale, $r=-.23, p=.026$) and the IRI subscale Empathic Concern in the first session ($r=-.26, p=.013$).

5.3.6 Age, Gender, and Education

Correlational analyses revealed several significantly positive linear associations between age and various empathy measures. Older participants tended to have higher levels of individual perspective taking (IRI subscale) in both the first ($r=.33, p=.001$) and the fourth session ($r=.34, p=.001$), individual empathic concern in the fourth session (IRI subscale; $r=.26, p=.014$), systemic perspective taking in the fourth session (CEEQ-GROUP subscale; $r=.25,$

$p=.019$), systemic empathic concern in the first (CEEQ-GROUP subscale; $r=.28$, $p=.005$) and in the fourth session ($r=.25$, $p=.016$), and systemic mirroring in the first session (CEEQ-GROUP subscale; $r=.23$, $p=.024$). Further, age was positively associated with the MET subscale Emotional Empathy in the first ($r=.29$, $p=.004$) and in the fourth session ($r=.32$, $p=.002$), as well as with the DCS group discussion parameter Empathic Statements in the first session ($r=.22$, $p=.030$). Age was not significantly correlated with an increase or decrease of any empathy measure between the first and the fourth session. Multiple regressions revealed a quadratic association between age and the IRI subscale Personal Distress in the first session ($\beta=-1.87$, $t=-2.02$, $p=.046$), indicating that middle-aged persons within the present sample had higher levels of personal distress than younger and older persons. Another quadratic association was found between age and an increase in scores on the IRI subscale Fantasy between the first and the fourth experimental session ($\beta=2.24$, $t=2.33$, $p=.022$), indicating that middle-aged persons differed less in individual fantasy between both sessions than younger and older participants.

Univariate ANOVAs revealed only few gender differences with regard to empathy measures. Women had significantly higher scores in individual empathic concern (IRI subscale) both in the first ($M_{\text{♂}}=3.36$, $SD_{\text{♂}}=.53$; $M_{\text{♀}}=3.57$, $SD_{\text{♀}}=.49$; $F(1,97)=4.17$; $p=.044$; $\eta^2=.03$) and in the fourth session ($M_{\text{♂}}=3.22$, $SD_{\text{♂}}=.64$; $M_{\text{♀}}=3.53$, $SD_{\text{♀}}=.50$; $F(1,90)=7.17$; $p=.009$; $\eta^2=.06$), as well as in individual personal distress (IRI subscale) in the first ($M_{\text{♂}}=2.55$, $SD_{\text{♂}}=.61$; $M_{\text{♀}}=2.91$, $SD_{\text{♀}}=.61$; $F(1,97)=7.90$; $p=.006$; $\eta^2=.07$) and in the fourth session ($M_{\text{♂}}=2.45$, $SD_{\text{♂}}=.64$; $M_{\text{♀}}=2.81$, $SD_{\text{♀}}=.56$; $F(1,90)=9.49$; $p=.003$; $\eta^2=.09$). With regard to systemic empathy measures, women scored significantly higher on the DCS group discussion parameter Empathic Feedback in the first ($M_{\text{♂}}=.25$, $SD_{\text{♂}}=.20$; $M_{\text{♀}}=.51$, $SD_{\text{♀}}=.35$; $F(1,97)=16.58$; $p<.001$; $\eta^2=.15$) and in the fourth session ($M_{\text{♂}}=.31$, $SD_{\text{♂}}=.32$; $M_{\text{♀}}=.47$, $SD_{\text{♀}}=.39$; $F(1,90)=4.27$; $p=.042$; $\eta^2=.04$). There were no further gender differences.

Years of education were significantly associated with the IRI subscale Perspective Taking in the first ($r=.32$, $p=.001$) and in the fourth session ($r=.26$, $p=.012$), with the IRI subscale Personal Distress in the fourth session ($r=.23$, $p=.031$), with the MET subscale Emotional Empathy in the fourth session ($r=.27$, $p=.010$), with the DCS group discussion parameter Empathic Feedback in the first session ($r=.22$, $p=.028$), and with an increase in scores on the MET subscale Emotional Empathy between the first and the fourth session ($r=.32$, $p=.002$).

5.4 Associations Between Individual and Systemic Empathy Measures

With regard to the first research question of the present investigation, partial correlations between all empathy measures were computed, including both points of measurement and the difference between both points of measurement, while controlling for all significantly associated control variables in each case, as reported in the previous section. Detailed tables with all computed partial correlations are presented in the Appendix (Section 8.3). To further examine the difference between systemic and individual empathy, three confirmatory factor analyses were computed for each point of measurement, based on the specific multitrait-multimethod model by Eid (2000; see Section 4.4.2 for details).

The results of the first factor analysis for the first point of measurement with the two latent higher-order factors Systemic and Individual Empathy can be derived from Figure 7. Bidirectional arrows represent correlations, which were specified between subscales of the same measure and the same subcomponent of empathy (cognitive vs. emotional empathy), whereas unidirectional arrows represent standardized factor loadings. Structural equation modelling with AMOS 22.0 revealed a significant difference between the observed and the expected covariance matrices of this specified model ($\chi^2(59)=84.23$; $p=.017$). Further, even if the computed Root Mean Square Error of Approximation indicated an acceptable model fit (RMSEA=.066; 90% CI=.029-.097), the Comparative Fit Index value clearly indicated a poor model fit (CFI=.865).

The second factor analysis was conducted for all cognitive empathy measures within the first experimental session (Figure 8). The observed and expected covariance matrices within this model did not significantly differ from each other ($\chi^2(4)=7.65$; $p=.11$), and the computed fit indices revealed an acceptable model fit (RMSEA=.097; 90% CI=.000-.201; CFI=.945). However, the standardized factor loading of the CEEQ-GROUP subscale Perspective Taking on the latent factor Systemic Cognitive Empathy was estimated with a value greater than one. Therefore, this factor loading represents a Heywood Case (Heywood, 1931; see also Brown, 2006, as well as Kolenikov and Bollen, 2012, for reviews on this phenomenon) and indicates an invalid factor solution.

With regard to the third model on emotional empathy measures within the first experimental session (Figure 9), the observed and expected covariance matrices did not significantly differ from each other ($\chi^2(4)=1.63$; $p=.80$) and all computed fit indices indicated a good model fit (RMSEA=.000; 90% CI=.000-.097; CFI=1.000).

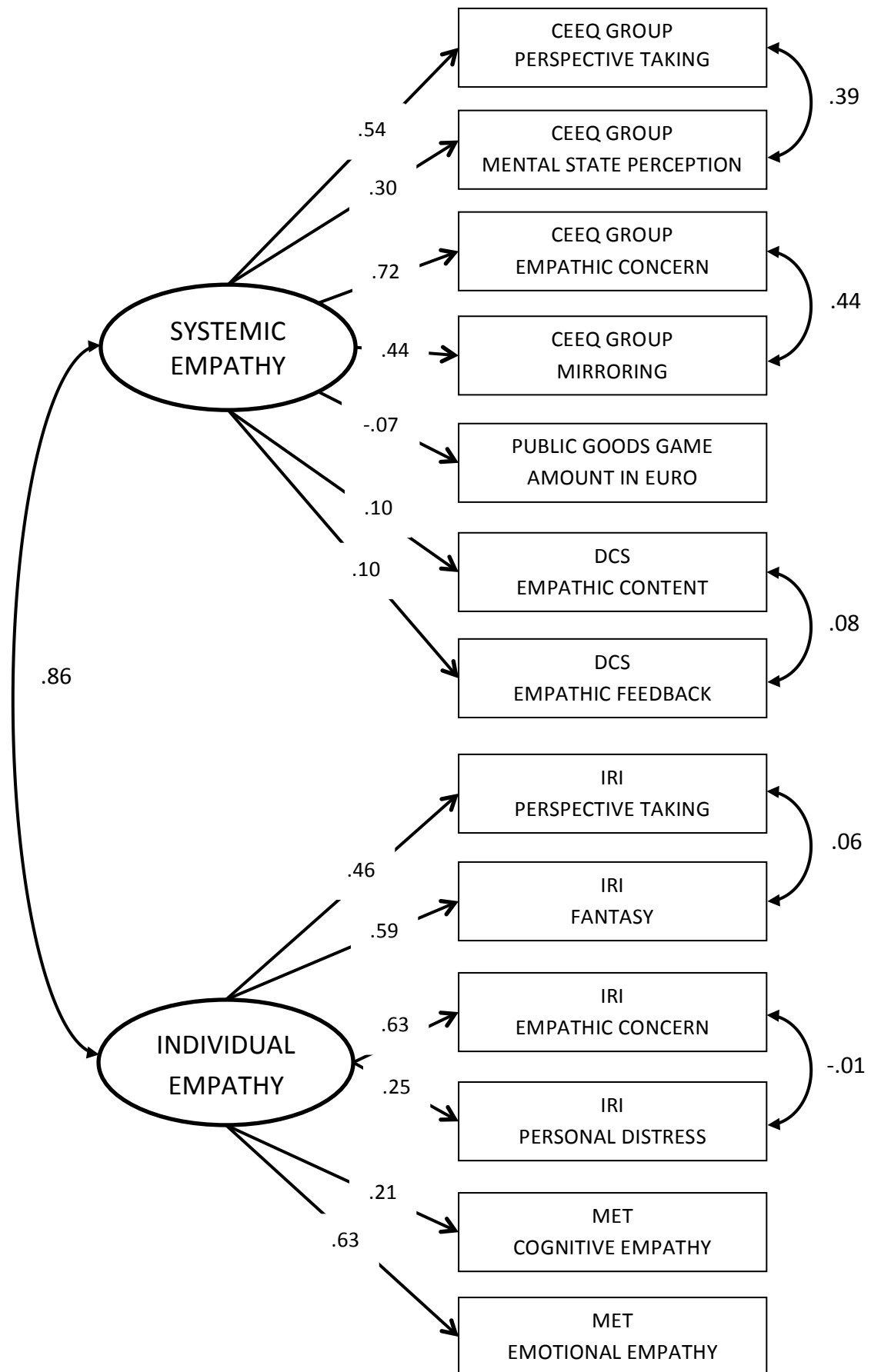


Figure 7. Factor Analysis of Empathy Measures in the First Experimental Session

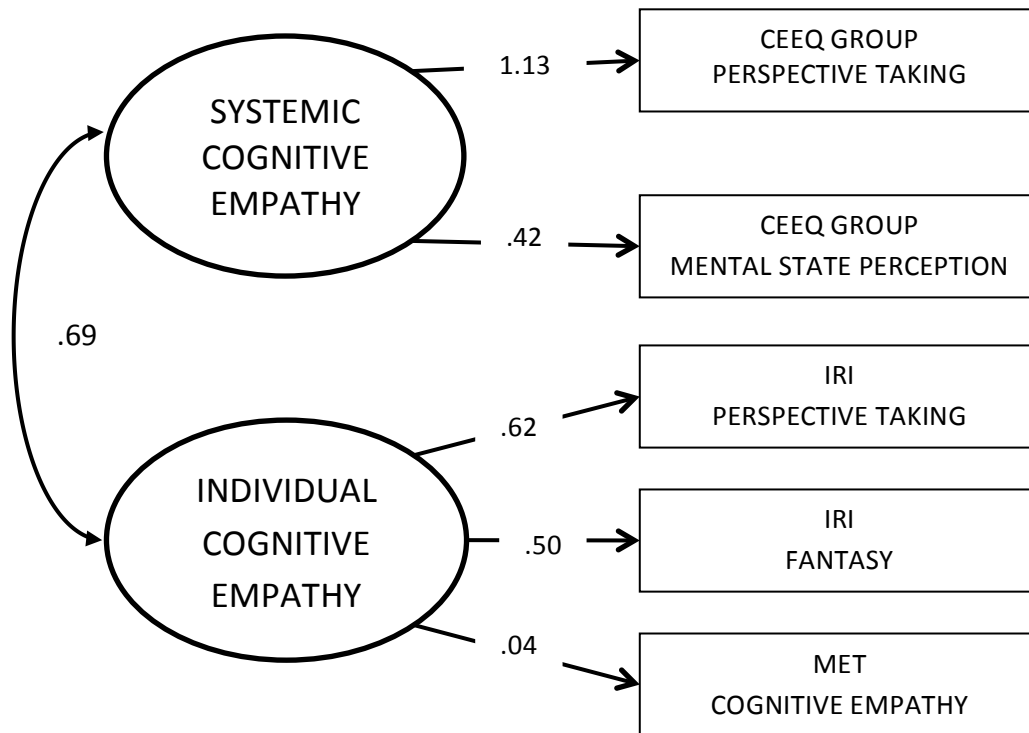


Figure 8. Factor Analysis of Cognitive Empathy Measures in the First Experimental Session

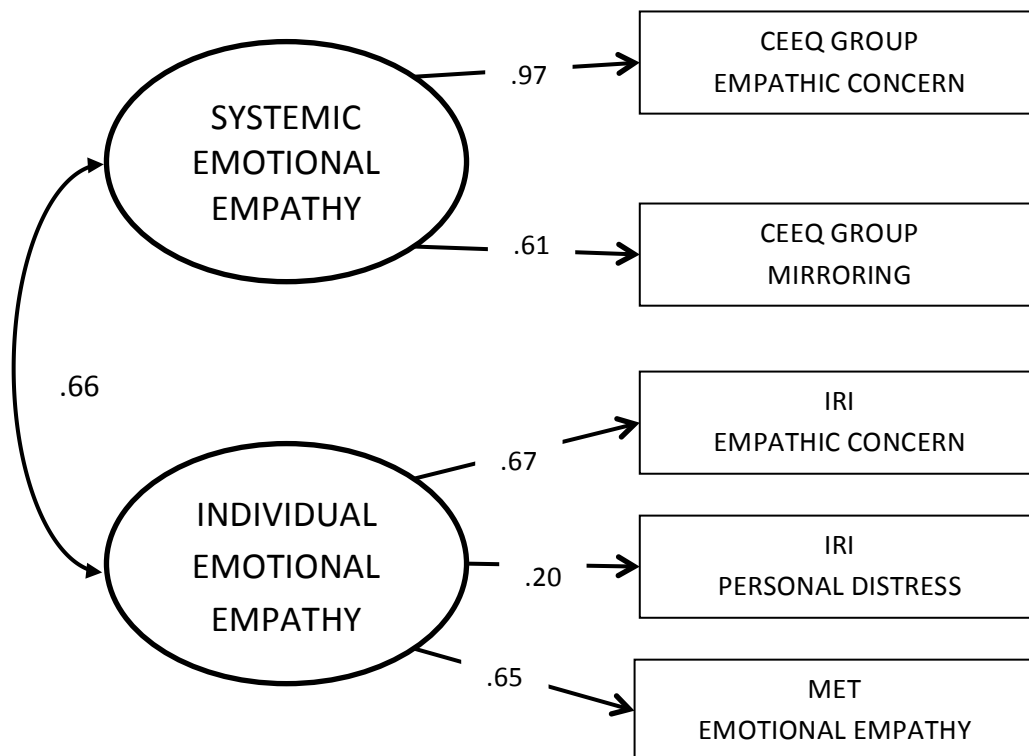


Figure 9. Factor Analysis of Emotional Empathy Measures in the First Experimental Session

The fourth factor analysis on all empathy measures within the fourth experimental session (Figure 10) revealed another Heywood Case, as the estimated correlation between the latent higher-order factors Systemic and Individual Empathy was greater than one. Further, the observed covariance matrices significantly differed from the expected covariance matrices ($\chi^2(59)=139.71$; $p<.001$) and both computed fit indices revealed a poor model fit (RMSEA=.12; 90% CI=.093-.14; CFI=.717).

With regard to the fifth model on cognitive empathy measures within the fourth experimental session (Figure 11), the difference between the observed and expected covariance matrices was marginally significant ($\chi^2(4)=8.63$; $p=.071$). The Root Mean Square Error of Approximation indicated a poor model fit (RMSEA=.11; 90% CI=.000-.21). On the other side, the Comparative Fit Index indicated an acceptable model fit (CFI=.943).

The final factor analysis on emotional empathy measures within the fourth experimental session (Figure 12) revealed also a marginally significant difference between the observed and the expected covariance matrices ($\chi^2(4)=9.04$; $p=.060$). Again, the Root Mean Square Error of Approximation indicated a poor model fit (RMSEA=.11; 90% CI=.000-.22). The Comparative Fit Index indicated a good model fit (CFI=.952).

In summary, only one of six computed factor analyses on the empathy measures within the present investigation clearly confirmed a two-factor solution with the latent higher-order factors Systemic and Individual Empathy. Anomalies as the reported Heywood Cases, indicators of poor model fit, and partially low factor loadings (especially with regard to the Public Goods Game) may be a result of increased common method variance within the remaining five models: If the covariance matrices between systemic and individual empathy measures of the same method (e.g., CEEQ-GROUP Perspective Taking and IRI Perspective Taking) are higher than the covariance matrices between systemic empathy measures of different methods (e.g., CEEQ-GROUP Perspective Taking and the DCS group discussion parameters), this would result in increased error variances within the computed factor models. Several partial correlations between empathy measures (see Section 8.3) support this explanation. However, alternative explanations as other measurement errors or the nonexistence of the latent constructs Systemic and Individual Empathy have to be taken into consideration as well (see Section 6.1 for a detailed discussion). Nonetheless, the following section proceeds with testing the main hypotheses of this investigation, notwithstanding the limited interpretability of the results of these analyses.

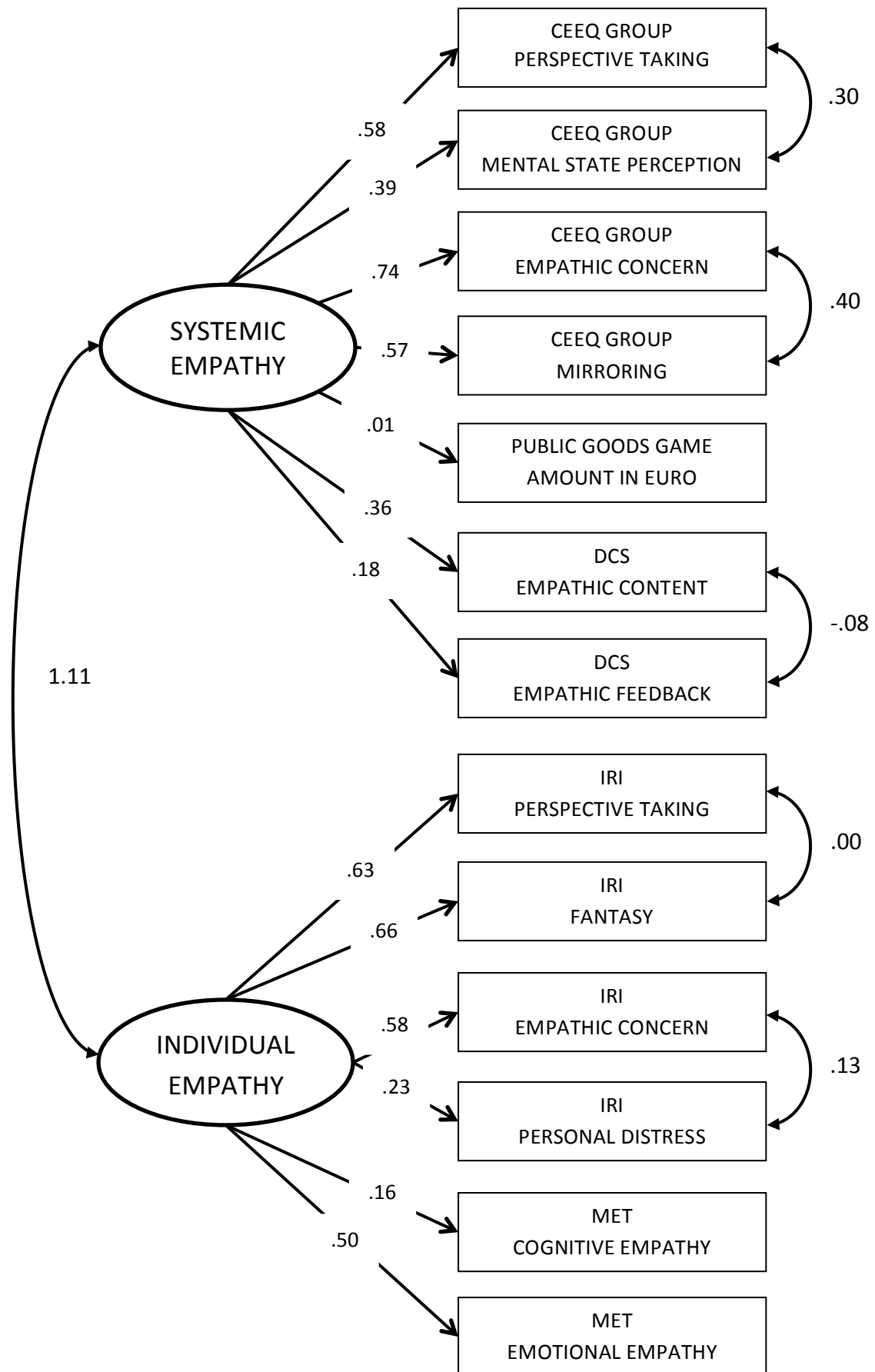


Figure 10. Factor Analysis of Empathy Measures in the Fourth Experimental Session

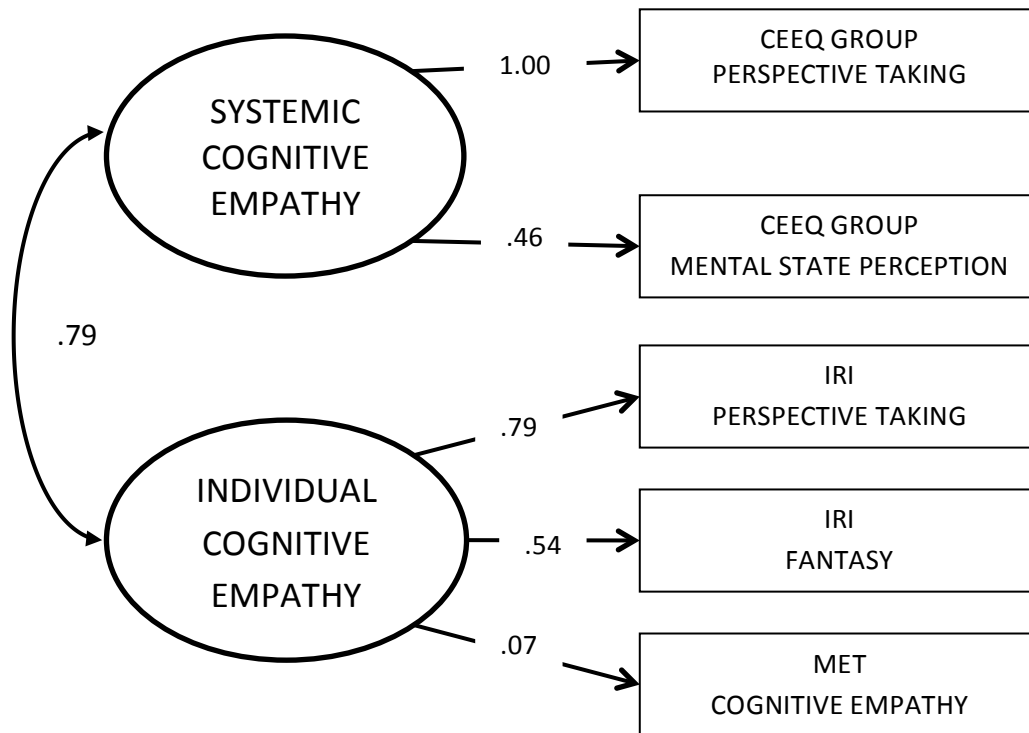


Figure 11. Factor Analysis of Cognitive Empathy Measures in the Fourth Experimental Session

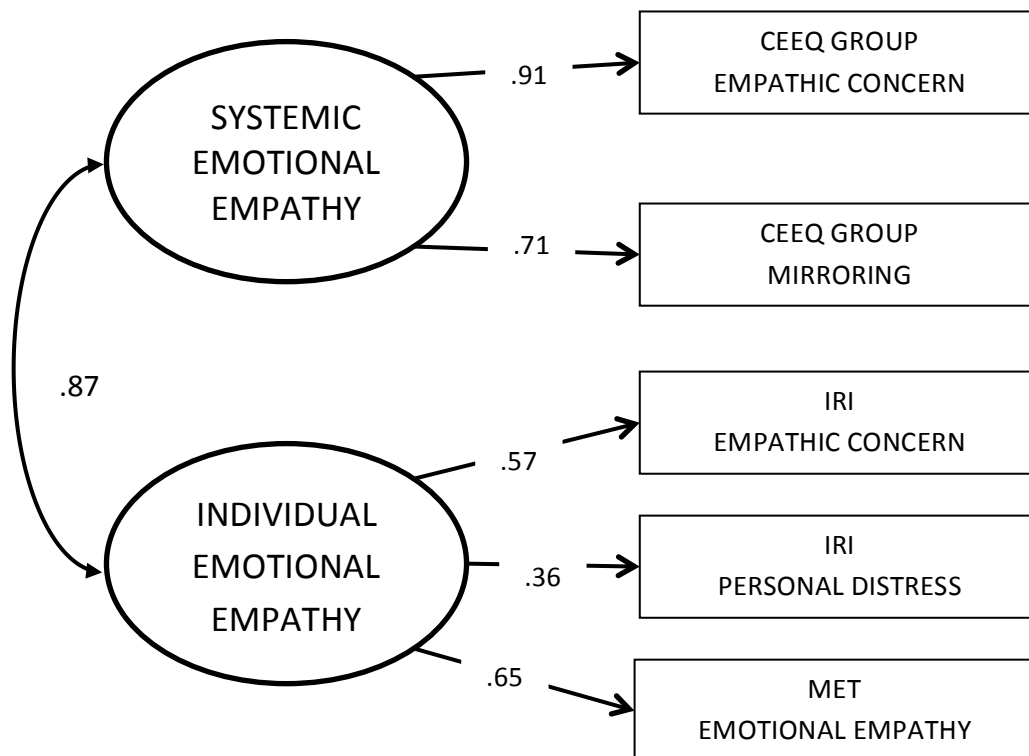


Figure 12. Factor Analysis of Emotional Empathy Measures in the Fourth Experimental Session

5.5 Impact of Group Parameters on Empathy

To investigate the impact of surface-level diversity and a shared goal on the development of individual empathy, analyses of covariance for repeated measures were computed, with experimental condition as the between-subjects factor, the difference between the first and second point of measurement as the within-subjects factor, the group affiliation of each participant as an additional random factor, and important control variables (see Section 5.3) as covariates. The impact of both group parameters on the development of systemic empathy was investigated by the same procedure and an additional multilevel model, considering the experimental condition as the first hierarchical level and the experimental group of each participant as the second hierarchical level of this analysis (see Section 4.4.3 for details). Considering the decreased power within the multilevel analyses, only the most important control variable, Social Desirability (see Section 5.1.2) was included within the multilevel model. The marginal means within all figures in this section are based on log-transformed scores (see Section 5.1.1 for raw descriptive statistics).

5.5.1 Questionnaires

IRI

The main effect of the experimental condition on the difference in scores on the IRI subscale Perspective Taking between the first and the second point of measurement was not significant ($F(1,90)=.03$; $p=.97$; $\eta^2=.00$; estimated marginal means of $M=1.50$ for the control condition, $M=1.55$ for the low diversity condition, $M=1.49$ for the shared goal condition at the first and the second point of measurement; see Figure 13). The null hypothesis of variance homogeneity between the three experimental conditions was not rejected by Levene's test ($p=.15$ for the first point of measurement, $p=.55$ for the second point of measurement). Social Desirability, Openness, Agreeableness, Conscientiousness, Group Identification, Group Climate, age, and education in years were included as covariates within this analysis. The analysis of covariance for repeated measures revealed no significant within-subjects or between-subjects effects of any of these covariates on the IRI subscale Perspective Taking.

There was also no significant effect of the experimental condition on score differences on the IRI subscale Fantasy ($F(1,90)=.05$; $p=.96$; $\eta^2=.00$; estimated marginal means of $M=1.45$ for the control condition, $M=1.46$ for the low diversity condition, $M=1.45$

for the shared goal condition at the first point of measurement, and $M=1.43$ for the control condition, $M=1.45$ for the low diversity condition, and $M=1.44$ for the shared goal condition at the second point of measurement; see Figure 14). Levene's test revealed homogenous variances within the three experimental conditions ($p=.13$ for the first point of measurement, $p=.11$ for the second point of measurement). With regard to the included covariates Social Desirability, Extraversion, Openness, Group Identification, the DCS group discussion parameter Affiliation, and squared age, there were no significant within-subjects or between-subjects effects.

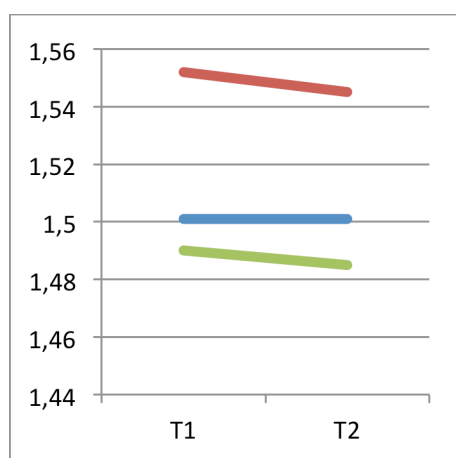


Figure 13. Marginal Means for the IRI Subscale Perspective Taking

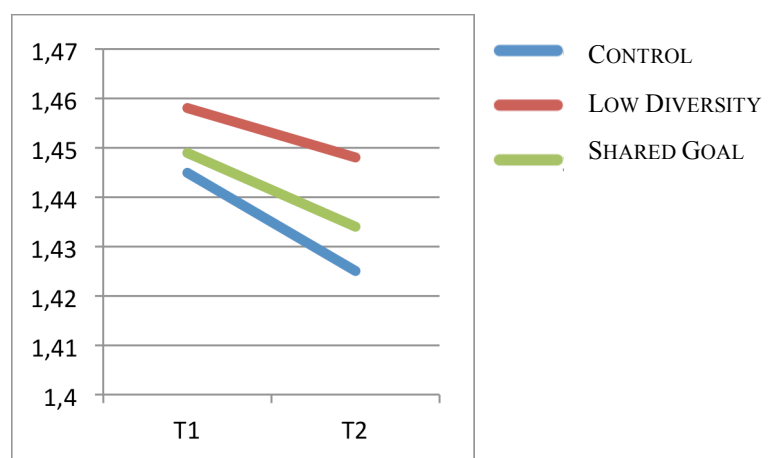


Figure 14. Marginal Means for the IRI Subscale Fantasy

The main effect of the experimental condition on the difference in scores on the IRI subscale Empathic Concern between the first and the second point of measurement was also not significant ($F(1,90)=.31$; $p=.73$; $\eta^2=.00$; estimated marginal means of $M=1.48$ for the control condition, $M=1.47$ for the low diversity condition, $M=1.52$ for the shared goal condition at the first point of measurement, and $M=1.46$ for the control condition, $M=1.45$ for the low diversity condition, and $M=1.52$ for the shared goal condition at the second point of measurement; see Figure 15). Levene's test indicated homogenous variances within all three experimental conditions ($p=.15$ for the first point of measurement, $p=.81$ for the second point of measurement). Social Desirability, Neuroticism, Extraversion, Openness, Agreeableness, Perceived Diversity, Diversity Beliefs, Goal Representation, the DCS group discussion parameters Dominance and Affiliation, Group Climate, and age were included as covariates within this analysis. There was a significant within-subjects effect of Goal

Representation at the first point of measurement on Empathic Concern ($F(1,90)=5.29$; $p=.025$; $\eta^2=.02$), as well as significant between-subjects effects of Neuroticism ($F(1,90)=5.53$; $p=.022$; $\eta^2=.11$), Agreeableness ($F(1,90)=7.035$; $p=.01$; $\eta^2=.14$), Perceived Diversity at the first point of measurement ($F(1,90)=4.30$; $p=.001$; $\eta^2=.09$), Diversity Beliefs at the first ($F(1,90)=11.03$; $p=.001$; $\eta^2=.22$) and the second point of measurement ($F(1,90)=13.48$; $p<.001$; $\eta^2=.27$), Goal Representation at the second point of measurement ($F(1,90)=6.34$; $p=.014$; $\eta^2=.13$), and Dominance at the first ($F(1,90)=7.55$; $p=.008$; $\eta^2=.15$) and the second point of measurement ($F(1,90)=5.06$; $p=.028$; $\eta^2=.10$).

The development of the scores of the IRI subscale Personal Distress did also not significantly differ among experimental conditions ($F(1,90)=1.79$; $p=.18$; $\eta^2=.01$; estimated marginal means of $M=1.34$ for the control condition, $M=1.32$ for the low diversity condition, $M=1.29$ for the shared goal condition at the first point of measurement, and $M=1.31$ for the control condition, $M=1.33$ for the low diversity condition, and $M=1.22$ for the shared goal condition at the second point of measurement; see Figure 16). The variances within the three experimental conditions were homogeneous ($p=.37$ for the first point of measurement, $p=.86$ for the second point of measurement). Social Desirability, Neuroticism, Group Identification, Group Climate, Perceived Diversity, Diversity Beliefs, squared age and education in years were included as covariates. The analysis revealed a significant within-subjects effect of Diversity Beliefs at the first ($F(1,90)=13.30$; $p<.001$; $\eta^2=.09$) and the second point of measurement on Personal Distress ($F(1,90)=8.21$; $p=.005$; $\eta^2=.05$), as well as significant between-subjects effects of Social Desirability ($F(1,90)=4.29$; $p=.042$; $\eta^2=.13$) and Neuroticism ($F(1,90)=18.33$; $p<.001$; $\eta^2=.55$).

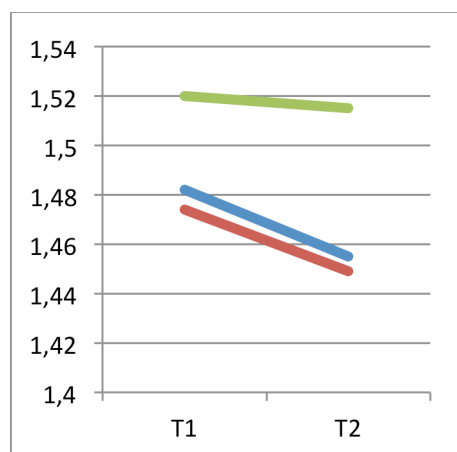


Figure 15. Marginal Means for the IRI Subscale Empathic Concern

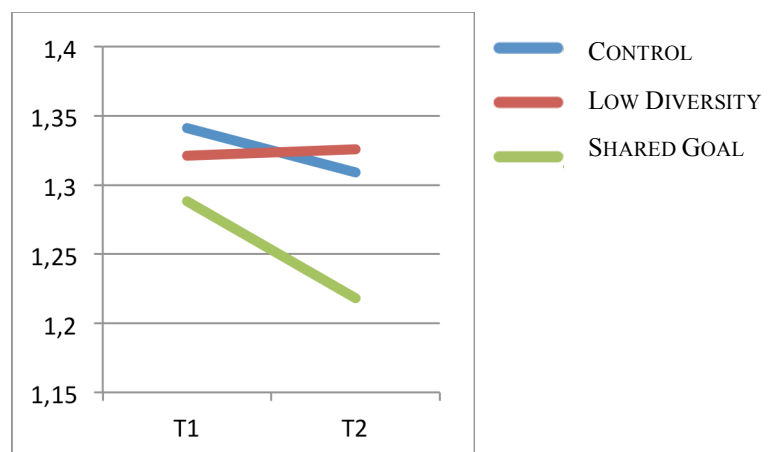


Figure 16. Marginal Means for the IRI Subscale Personal Distress

CEEQ-GROUP

With regard to the development of the scores of the CEEQ-GROUP subscale Perspective Taking between the first and the second point of measurement, the analysis of covariance for repeated measures revealed no significant difference between the three experimental conditions ($F(1,90)=.09$; $p=.92$; $\eta^2=.00$; estimated marginal means of $M=1.44$ for the control condition, $M=1.48$ for the low diversity condition, $M=1.42$ for the shared goal condition at the first point of measurement, and $M=1.45$ for the control condition, $M=1.50$ for the low diversity condition, and $M=1.44$ for the shared goal condition at the second point of measurement; see Figure 17). The variances within the three experimental conditions were homogeneous ($p=.88$ for the first point of measurement, $p=.44$ for the second point of measurement). Out of the included covariates Social Desirability, Agreeableness, Conscientiousness, Group Climate, the DCS group discussion parameter Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation, and age, the ANCOVA revealed significant within-subjects effects of Group Climate in the third experimental session ($F(1,90)=4.04$; $p=.048$; $\eta^2=.02$), the DCS group discussion parameter Affiliation in the fourth experimental session ($F(1,90)=6.48$; $p=.013$; $\eta^2=.03$), and Perceived Diversity in the first experimental session ($F(1,90)=5.90$; $p=.018$; $\eta^2=.03$) on the CEEQ-GROUP subscale Perspective Taking. The analysis did not reveal a significant between-subjects effect of any included covariate.

The additional multilevel-analysis revealed also no significant difference between the three experimental conditions with regard to the change in scores of the CEEQ-GROUP subscale Perspective Taking ($F(2)=.99$; $p=.40$).

Similarly, the score difference on the CEEQ-GROUP subscale Mental State Perception did not significantly differ between the three experimental conditions ($F(1,90)=1.35$; $p=.27$; $\eta^2=.01$; estimated marginal means of $M=1.44$ for the control condition, $M=1.42$ for the low diversity condition, $M=1.41$ for the shared goal condition at the first point of measurement, and $M=1.41$ for the control condition, $M=1.43$ for the low diversity condition, and $M=1.44$ for the shared goal condition at the second point of measurement; see Figure 18). The null hypothesis of homogeneous variances within the three experimental conditions was not rejected by Levene's test ($p=.22$ for the first point of measurement, $p=.80$ for the second point of measurement). Social Desirability, Agreeableness, Group Identification, Group Climate, and the DCS group discussion parameter Affiliation were included as covariates in this analysis. There were significant within-subjects effects of Social Desirability

($F(1,90)=4.67$; $p=.034$; $\eta^2=.03$), Agreeableness ($F(1,90)=8.92$; $p=.004$; $\eta^2=.06$), and Affiliation at the first point of measurement ($F(1,90)=10.97$; $p=.001$; $\eta^2=.07$) on the CEEQ-GROUP subscale Mental State Perception, but no significant between-subjects effects.

Again, the additional multilevel model revealed no significant difference between the three experimental conditions with regard to the development of the CEEQ-GROUP subscale Mental State Perception between the first and the fourth experimental session ($F(2)=1.00$; $p=.37$).

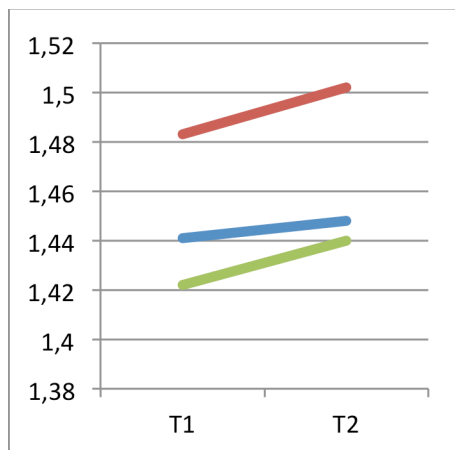


Figure 17. Marginal Means for the CEEQ Subscale Perspective Taking

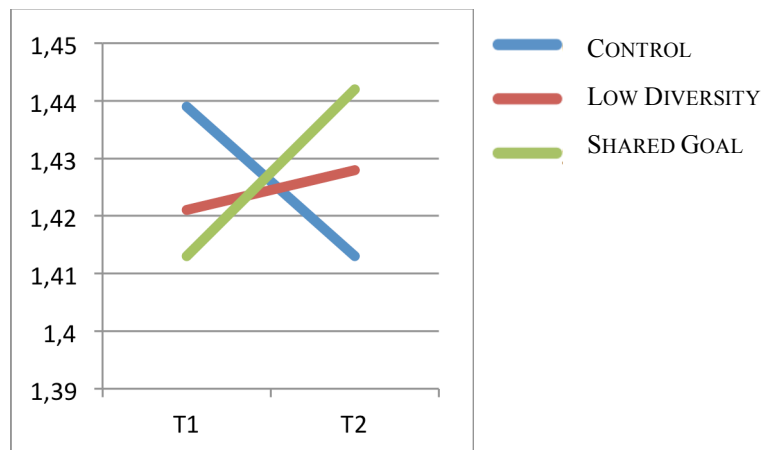


Figure 18. Marginal Means for the CEEQ Subscale Mental State Perception

The development of the CEEQ-GROUP subscale Empathic Concern did not significantly differ between the three experimental conditions, too ($F(1,90)=2.24$; $p=.11$; $\eta^2=.01$; estimated marginal means of $M=1.46$ for the control condition, $M=1.54$ for the low diversity condition, $M=1.49$ for the shared goal condition at the first point of measurement, and $M=1.48$ for the control condition, $M=1.50$ for the low diversity condition, and $M=1.52$ for the shared goal condition at the second point of measurement; see Figure 19). The variances within the three experimental conditions were homogeneous ($p=.72$ for the first point of measurement, $p=.99$ for the second point of measurement). Out of the included covariates Social Desirability, Agreeableness, Conscientiousness, Group Identification, Perceived Diversity, the DCS group discussion parameter Affiliation, Goal Representation, and age, there were significant within-subjects effects of Group Identification in the third experimental session ($F(1,90)=4.42$; $p=.039$; $\eta^2=.03$), Perceived Diversity in the fourth experimental session ($F(1,90)=5.18$; $p=.026$; $\eta^2=.03$), and Goal Representation in the first experimental session ($F(1,90)=7.77$; $p=.007$; $\eta^2=.05$) on the changes in systemic empathic

concern. Further, the analysis revealed significant between-subjects effects of the covariates Group Identification in the first experimental session ($F(1,90)=4.53$; $p=.037$; $\eta^2=.11$), Group Identification in the fourth experimental session ($F(1,90)=5.49$; $p=.022$; $\eta^2=.13$), and age ($F(1,90)=4.05$; $p=.048$; $\eta^2=.10$).

The multilevel analysis revealed also no significant difference in the development of systemic empathic concern between the three experimental conditions ($F(2)=.47$; $p=.63$).

There was a significant difference between the three experimental conditions in score development of the CEEQ-GROUP subscale Mirroring ($F(1,90)=3.85$; $p=.026$; $\eta^2=.03$). However, an increase of the marginal means in Mirroring within the control condition was associated with a decrease of the marginal means in Mirroring within both experimental conditions, which is contrary to both hypotheses (estimated marginal means of $M=1.34$ for the control condition, $M=1.38$ for the low diversity condition, $M=1.35$ for the shared goal condition at the first point of measurement, and $M=1.40$ for the control condition, $M=1.34$ for the low diversity condition, and $M=1.34$ for the shared goal condition at the second point of measurement; see Figure 20). The variances within the experimental conditions were homogeneous across both points of measurement ($p=.59$ in the first experimental session, $p=.50$ in the fourth experimental session). Social Desirability, Extraversion, Group Identification, the DCS group discussion parameter Affiliation, Perceived Diversity, and age were included as covariates. There was one significant within-subjects effect of Group Identification in the third experimental session on differences in Mirroring ($F(1,90)=4.62$; $p=.035$; $\eta^2=.03$), and no significant between-subjects effects.

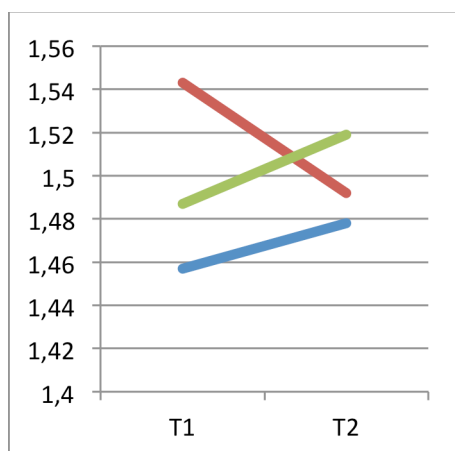


Figure 19. Marginal Means for the CEEQ Subscale Empathic Concern

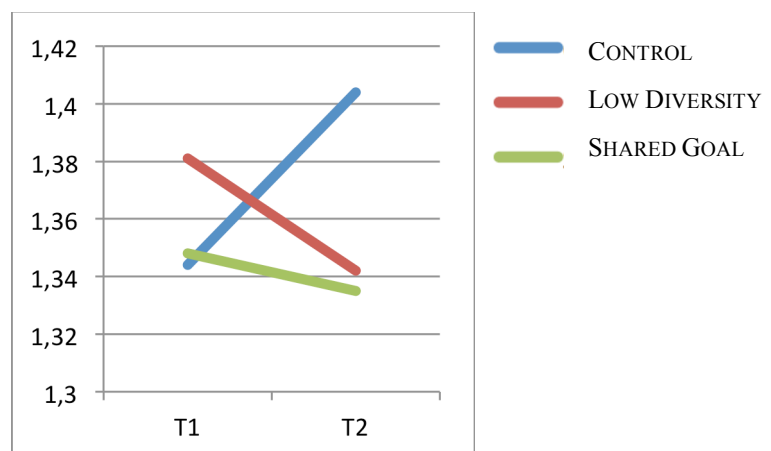


Figure 20. Marginal Means for the CEEQ Subscale Mirroring

However, the additional multilevel model did not confirm the significant effect for systemic mirroring of the ANCOVA for repeated measures. The score development in systemic mirroring did not significantly differ between the three experimental conditions in this analysis ($F(2)=.91$; $p=.42$), indicating the possibility that the previously reported significant effect was caused by random group-level variance.

In summary, the not significant results on the individual empathy measure IRI are in line with the hypotheses of the present investigation, while the results on the systemic empathy measure CEEQ-GROUP are not in line with the two established hypotheses.

5.5.2 Multifaceted Empathy Test

An ANCOVA for repeated measures revealed no significant difference between the three experimental conditions with regard to changes in the MET subscale Cognitive Empathy between the first and the fourth experimental session ($F(1,90)=.31$; $p=.73$; $\eta^2=.00$; estimated marginal means of $M=.38$ for the control condition, $M=.37$ for the low diversity condition, $M=.44$ for the shared goal condition at the first point of measurement, and $M=.38$ for the control condition, $M=.38$ for the low diversity condition, and $M=.46$ for the shared goal condition at the second point of measurement; see Figure 21). The null hypothesis of variance homogeneity between the three experimental conditions was not rejected by Levene's test ($p=.71$ for the first experimental session; $p=.35$ for the fourth experimental session). With regard to the included covariates Social Desirability, Extraversion, Openness, Group Climate, the DCS group discussion parameters Affiliation and Dominance, and Goal Representation, no significant within-subjects effects on changes in scores of the MET subscale Cognitive Empathy were revealed. However, there were significant between-subjects effects of Social Desirability ($F(1,90)=4.67$; $p=.034$; $\eta^2=.07$), Extraversion ($F(1,90)=5.48$; $p=.022$; $\eta^2=.09$), Dominance in the first experimental session ($F(1,90)=5.09$; $p=.027$; $\eta^2=.08$), Affiliation in the first experimental session ($F(1,90)=4.75$; $p=.033$; $\eta^2=.08$), and Goal Representation in the fourth experimental session ($F(1,90)=4.62$; $p=.035$; $\eta^2=.07$) on the MET subscale Cognitive Empathy.

The score development of the MET subscale Emotional Empathy did also not significantly differ between the three experimental conditions ($F(1,90)=.48$; $p=.62$; $\eta^2=.01$; estimated marginal means of $M=1.72$ for the control condition, $M=1.75$ for the low diversity

condition, $M=1.77$ for the shared goal condition at the first point of measurement, and $M=1.73$ for the control condition, $M=1.79$ for the low diversity condition, and $M=1.75$ for the shared goal condition at the second point of measurement; see Figure 22). The variances within the three experimental conditions were homogeneous ($p=.68$ for the first experimental session; $p=.98$ for the fourth experimental session). Social Desirability, Openness, Agreeableness, Conscientiousness, Group Identification, the DCS group discussion parameter Affiliation, Goal Representation, age, and education in years were included as covariates within this analysis. The ANCOVA for repeated measures revealed significant within-subjects effects of Goal Representation in the first experimental session ($F(1,90)=8.27$; $p=.005$; $\eta^2=.10$) and education in years ($F(1,90)=5.21$; $p=.025$; $\eta^2=.06$), as well as a between-subjects effect of Social Desirability ($F(1,90)=4.73$; $p=.033$; $\eta^2=.39$) on the MET subscale Emotional Empathy.

The not significant findings on both MET subscales are in line with the hypotheses of the present investigation.

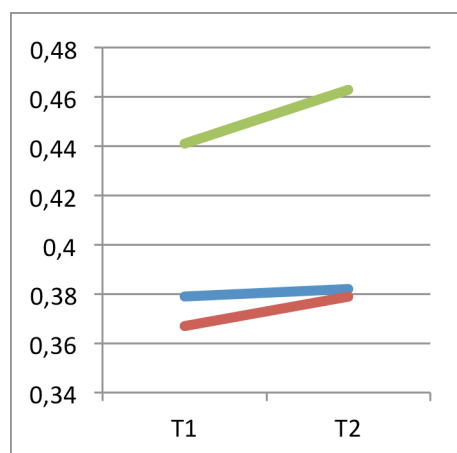


Figure 21. Marginal Means for the MET Subscale Cognitive Empathy

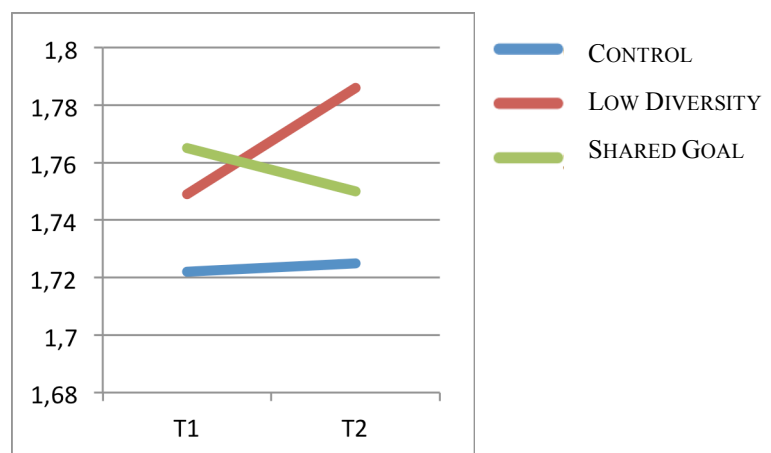


Figure 22. Marginal Means for the MET Subscale Emotional Empathy

5.5.3 Public Goods Game

An ANCOVA for repeated measures revealed that the change in the amount of money that was paid into the common cash box between the first and fourth experimental session did not significantly differ between the three experimental conditions ($F(1,90)=1.36$; $p=.27$; $\eta^2=.18$; estimated marginal means of $M=1.29$ for the control condition, $M=1.18$ for the low diversity condition, $M=1.37$ for the shared goal condition at the first point of measurement,

and $M=1.16$ for the control condition, $M=1.09$ for the low diversity condition, and $M=1.55$ for the shared goal condition at the second point of measurement; see Figure 23). Levene's test revealed that the variances within the three experimental conditions were homogeneous ($p=.66$ for the first experimental session; $p=.27$ for the fourth experimental session). Out of the included covariates Social Desirability, Neuroticism, Conscientiousness, Group Identification, Group Climate, Perceived Diversity, the DCS group discussion parameters Dominance and Affiliation, and Goal Representation, no significant within-subjects effects on the results of the Public Goods Game were found. Significant between-subjects effects on the Public Goods Game were revealed for Conscientiousness ($F(1,90)=6.50$; $p=.013$; $\eta^2=1.67$), Group Identification in the first experimental session ($F(1,90)=8.58$; $p=.005$; $\eta^2=2.20$), and Affiliation in the first experimental session ($F(1,90)=8.53$; $p=.005$; $\eta^2=2.19$).

However, after considering the specific group-level variance within an additional multilevel model, a significant difference between the three experimental conditions with regard to changes in the results of the Public Goods Game between the first and the fourth experimental session was revealed ($F(2)=4.03$; $p=.039$). The marginal means from the first analysis (Figure 23) indicate a predominant increase within the shared goal condition. Considering the decreased power-level of the multilevel analysis (see Section 4.4.3), the probability that this significant finding is a random effect may be rated as low. As this finding is moreover the only significant effect being at least partially in line with the established hypotheses, it is further examined in Section 5.6.

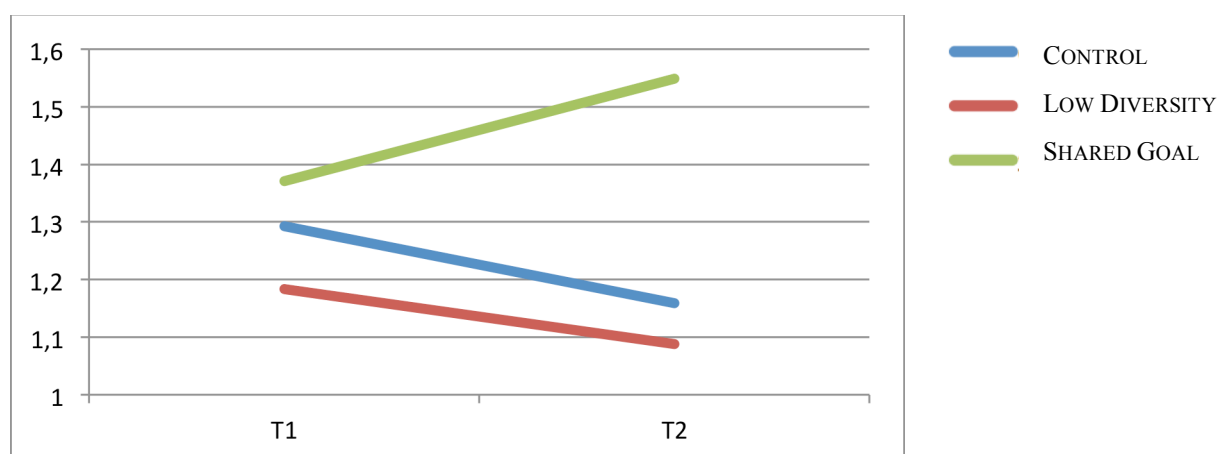


Figure 23. Marginal Means for the Results of the Public Goods Game

5.5.4 Group Discussion Parameters

The differences in the DCS group discussion parameter Empathic Content between the first and the fourth experimental session did not significantly differ between the three experimental conditions ($F(1,90)=1.64$; $p=.20$; $\eta^2=.03$; estimated marginal means of $M=.14$ for the control condition, $M=.12$ for the low diversity condition, $M=.15$ for the shared goal condition at the first point of measurement, and $M=.10$ for the control condition, $M=.17$ for the low diversity condition, and $M=.10$ for the shared goal condition at the second point of measurement; see Figure 24). The null hypothesis of variance homogeneity between the three experimental conditions was not rejected ($p=.73$ for the first experimental session; $p=.20$ for the fourth experimental session). Social Desirability, Openness, Conscientiousness, the DCS group discussion parameters Affiliation and Dominance, and Perceived Diversity were included as covariates. There was a significant within-subjects effect of Dominance in the first experimental session ($F(1,90)=4.21$; $p=.043$; $\eta^2=.07$) and Affiliation in the fourth experimental session ($F(1,90)=8.23$; $p=.005$; $\eta^2=.14$) on the DCS group discussion parameter Empathic Content, as well as significant between-subjects effects of Openness ($F(1,90)=11.57$; $p=.001$; $\eta^2=.26$) and Dominance in the fourth experimental session ($F(1,90)=5.02$; $p=.028$; $\eta^2=.11$).

The multilevel model revealed also no significant difference between the three experimental conditions with regard to changes in the scores of Empathic Content between both points of measurement ($F(2)=2.01$; $p=.14$).

Score changes in the DCS group discussion parameter Empathic Feedback between the first and the fourth experimental session did also not differ between the three experimental conditions. However, there was a trend towards statistical significance ($F(1,90)=2.85$; $p=.064$; $\eta^2=.07$). The estimated marginal means of $M=.29$ for the control condition, $M=.35$ for the low diversity condition, $M=.29$ for the shared goal condition at the first point of measurement, and $M=.37$ for the control condition, $M=.32$ for the low diversity condition, and $M=.24$ for the shared goal condition at the second point of measurement indicate an effect contrary to both established hypotheses (see Figure 25). The variances between the three experimental conditions were homogeneous ($p=.40$ for the first experimental session; $p=.66$ for the fourth experimental session). With regard to the included covariates Social Desirability, Agreeableness, Group Identification, and education in years, the ANCOVA for repeated measures revealed a significant within-subjects effect of

Social Desirability on Empathic Feedback ($F(1,90)=4.41$; $p=.039$; $\eta^2=.10$) and a significant between-subjects effect of Agreeableness ($F(1,90)=4.87$; $p=.030$; $\eta^2=.40$).

The additional multilevel analysis did not confirm the initial statistical trend: the model revealed no significant differences between the three experimental conditions with regard to changes in Empathic Feedback between the first and the fourth experimental session ($F(2)=.63$; $p=.54$). This indicates the possibility that the initial statistical trend was affected by random group-level variance, which was not considered within the first analysis of covariance on individual level.

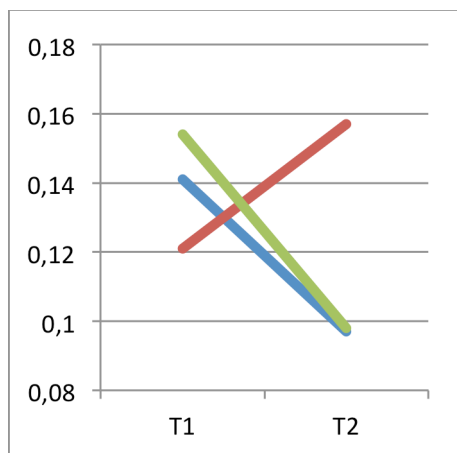


Figure 24. Marginal Means for the DCS Parameter Empathic Content

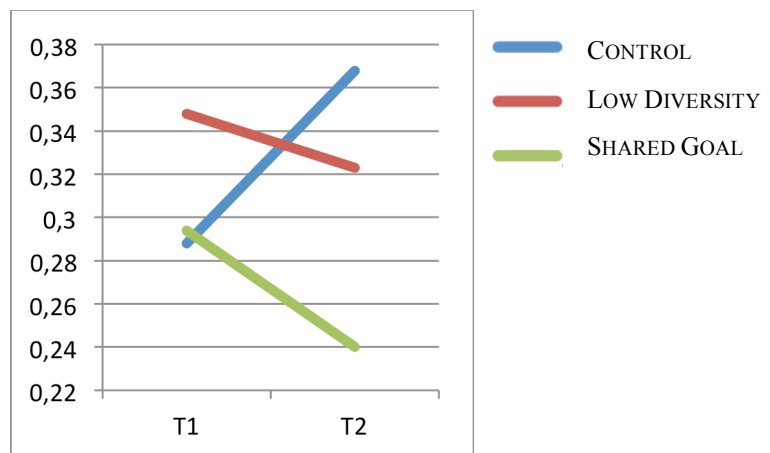


Figure 25. Marginal Means for the DCS Parameter Empathic Feedback

Summing up, the results on the DCS group discussion parameters did not confirm both established hypotheses.

5.6 Follow-Up Analyses on Possible Moderator and Mediator Effects

The multilevel analyses in Section 5.5 revealed a significant effect of the experimental condition on changes in the results of the Public Goods Game between the first and the fourth experimental session. As the univariate ANOVAs in Section 5.1.2 revealed a significant difference between the three experimental conditions with regard to changes in group climate between the first and the fourth experimental session, and with regard to static group climate in the fourth experimental session, a possible moderating or mediating effect

of dynamic or static group climate on the interrelationship between group parameters and the results of the Public Goods Game has to be taken into consideration.

In a first step towards investigating these possibilities, a moderated regression with dummy-coded experimental conditions, the most influential covariate social desirability (see Section 5.1.2), changes in group climate, and the products of dummy-coded experimental conditions and changes in group climate as predictors, and changes in the results of the Public Goods Game as dependent variable were conducted. This regression revealed no significant moderating effect of changes in group climate on the interrelationship between experimental condition and the results of the Public Goods Game (see Table 12).

Table 12

Moderated Regression Testing for a Moderating Effect of Changes in Group Climate on the Interrelationship Between Experimental Condition and Changes in the Results of the Public Goods Game

Predictor	Coefficients				
	B	SE	β	t	p
Dummy: Low Diversity Condition	.41	.37	.14	1.11	.27
Dummy: Shared Goal Condition	1.01	.40	.35	2.51	.014
Social Desirability	.83	1.24	.07	.67	.50
Δ Group Climate	.18	.19	.12	.96	.34
Low Diversity Condition * Δ Group Climate	.23	.50	.05	.46	.65
Shared Goal Condition * Δ Group Climate	-.35	.45	-.10	-.79	.44

Note. Dependent variable was the Δ between the results of the Public Goods Game in the first and in the fourth experimental session. Adjusted r^2 = .114. N = 91.

Another moderated regression considering static group climate in the fourth experimental session as an alternative predictor did also not reveal a significant moderating effect of static group climate on the interrelationship between experimental condition and changes in the results of the Public Goods Game (low diversity condition * group climate in the fourth session: β = .00, t = .00, p = 1.00; shared goal condition * group climate in the fourth session: β = .31, t = .19, p = .85). A third moderated regression with the same predictors as the first regression (see Table 12), but static results of the Public Goods Game in the fourth experimental session as dependent variable did also not reveal a significant moderating

effect of changes in group climate on the interrelationship between experimental condition and the Public Goods Game (low diversity condition * group climate in the fourth session: $\beta = -.01$, $t = -.07$, $p = .94$; shared goal condition * group climate in the fourth session: $\beta = -.03$, $t = -.22$, $p = .83$). Finally, a fourth moderated regression with static group climate in the fourth experimental session as an alternative predictor and static results of the Public Goods Game in the fourth experimental session as the dependent variable was conducted. Even if this regression did not confirm a moderating effect of group climate on the interrelationship between experimental condition and the Public Goods Game, it pointed to a possible mediator effect: the effect of static group climate on the results of the Public Goods Game in the fourth experimental session was significant, while the significant effect of the experimental condition disappeared (see Table 13).

Table 13

Moderated Regression Testing for a Moderating Effect of Static Group Climate on the Interrelationship Between Experimental Condition and Static Results of the Public Goods Game in the Fourth Experimental Session

Predictor	Coefficients				
	B	SE	β	t	p
Dummy: Low Diversity Condition	.53	1.27	.47	.42	.68
Dummy: Shared Goal Condition	-1.46	1.82	-1.27	-.80	.43
Social Desirability	-.10	.49	-.02	-.21	.84
Group Climate Fourth Session	.58	.26	.26	2.24	.028
Low Diversity Condition * Group Climate	-.29	.78	-.41	-.37	.72
Shared Goal Condition * Group Climate	1.01	1.10	1.48	.92	.36

Note. Dependent variable was the result of the Public Goods Game in the fourth experimental session. Adjusted $r^2 = .141$. $N = 91$.

A subsequent mediation analysis confirmed a mediating effect of static group climate on the interrelationship between the shared goal condition and static results of the Public Goods Game. A regression on the results of the Public Goods Game in the fourth experimental session with the three predictors dummy coded low diversity condition, dummy coded shared goal condition, and social desirability revealed a significant effect of the shared goal condition on the Public Goods Game ($\beta = .30$, $t = 2.43$, $p = .017$), but not of the

low diversity condition ($\beta = -.01$, $t = -.09$, $p = .93$), which reflects the results of the multilevel analysis reported in Section 5.5.3. Adding static group climate in the fourth experimental session as an additional predictor within this regression eliminated the significant effect of dummy coded shared goal condition on the Public Goods Game ($\beta = .20$, $t = 1.59$, $p = .12$), while the added predictor significantly affected the dependent variable ($\beta = .27$, $t = 2.54$, $p = .013$). As another regression analysis revealed a significant effect of the shared goal condition on static group climate in the fourth experimental session ($\beta = .36$, $t = 3.13$, $p = .002$), all requirements of a classical mediator effect are met. The mediation is summarized in Figure 26. In the first experimental session, there was neither a significant effect of the shared goal condition ($\beta = .00$, $t = -.01$, $p = 1$) nor of group climate ($\beta = .09$, $t = .82$, $p = .41$) on the results of the Public Goods Game. All presented analyses represent strong arguments for the possibility, that the significant multilevel effect reported in Section 5.5.3 might be at least partially mediated by group climate.

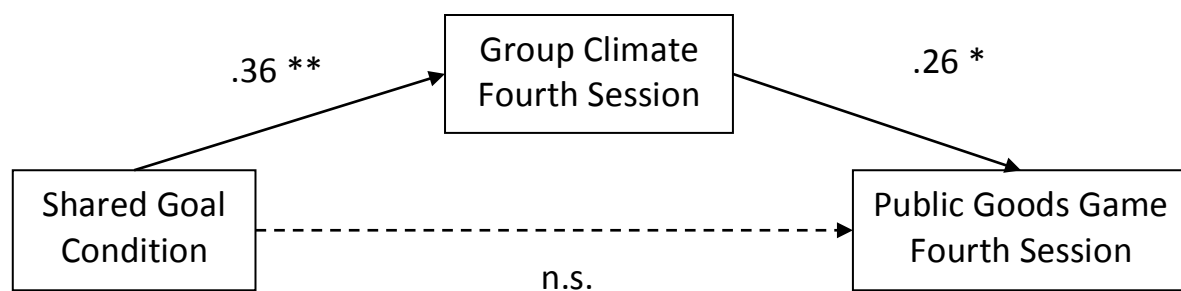


Figure 26. Mediator Effect of Static Group Climate on the Interrelationship Between the Shared Goal Condition and the Effects of the Public Goods Game Within the Fourth Session.

Further follow-up analyses were conducted to investigate a possible moderating or mediating effect of group identification on the interrelationship between the experimental condition and the results of the Public Goods Game, which would be in line with previous investigations on the role of group identification for empathy-motivated helping behavior (Simon et al., 2000; see Section 4.3.1.5). A moderated regression with dummy-coded experimental conditions, social desirability, changes in group identification, and the products of dummy-coded experimental conditions and changes in group identification as predictors, and changes in the results of the Public Goods Game as dependent variable revealed a marginally significant moderator effect for the shared goal condition (Table 14): The difference in the results of the Public Goods Game within the shared goal condition between

the first and the fourth experimental session seems to be moderated by concurrent changes in group identification ($\beta = -.24$, $t = -1.95$, $p = .055$).

Table 14

Moderated Regression Testing for a Moderating Effect of Changes in Group Identification on the Interrelationship Between Experimental Condition and Changes in the Results of the Public Goods Game

Predictor	Coefficients				
	B	SE	β	t	p
Dummy: Low Diversity Condition	.50	.35	.17	1.40	.17
Dummy: Shared Goal Condition	1.10	.35	.38	3.14	.002
Social Desirability	.97	1.23	.09	.79	.43
Δ Group Identification	.14	.23	.09	.58	.56
Low Diversity Condition * Δ Group Identification	.14	.37	.05	.38	.71
Shared Goal Condition * Δ Group Identification	-.82	.42	-.24	-1.95	.055

Note. Dependent variable was the Δ between the results of the Public Goods Game in the first and in the fourth experimental session. Adjusted $r^2 = .146$. $N = 91$.

Static group identification in the fourth experimental session did not significantly moderate the interrelationship between experimental condition and the results of the Public Goods Game (low diversity condition * group identification in the fourth session: $\beta = .62$, $t = .94$, $p = .35$; shared goal condition * group identification in the fourth session: $\beta = -.67$, $t = -.94$, $p = .35$). Moderated regressions did also not reveal a significant moderating effect of changes in group identification (low diversity condition * changes in group identification: $\beta = .00$, $t = .02$, $p = .99$; shared goal condition * changes in group identification: $\beta = -.17$, $t = -1.36$, $p = .18$) or static group identification (low diversity condition * group identification in the fourth session: $\beta = -.09$, $t = -.13$, $p = .89$; shared goal condition * group identification in the fourth session: $\beta = -1.06$, $t = -1.49$, $p = .14$) on the interrelationship between experimental condition and static results of the Public Goods Game in the fourth experimental session. Within the last regression analysis on static group identification, the level of group identification in the fourth session was a significant predictor for the results of the Public Goods Game within this session ($\beta = .32$, $t = 2.28$, $p = .025$). However, as the effect of the shared goal condition still remained significant within a subsequent mediation analysis ($\beta = .26$,

$t=2.17, p=.033$), the requirements for a mediator effect are not met in this case. Summing up, the presented regression analyses point to the possibility that the effect reported in Section 5.5.3 might be moderated by changes in group identification, too.

Finally, after investigating the role of group climate and group identification, further multiple regressions were conducted to spotlight the function of another potentially influential variable: Perceived Diversity. Even if prior analyses revealed no significant effect of experimentally manipulated surface-level diversity on any systemic empathy measure (see Section 5.1 and 5.4), perceived surface-level diversity might be more relevant for changes in systemic empathy than objective surface-level diversity (see Section 2.4.4.1 and 4.3.1.6). Interestingly, a moderated regression with dummy-coded experimental conditions, social desirability, changes in perceived diversity, and the products of dummy-coded experimental conditions and changes in perceived diversity as predictors, and changes in the results of the Public Goods Game as dependent variable revealed another marginally significant moderator effect of perceived diversity within the low diversity condition, but not within the shared goal condition (Table 15).

Table 15

Moderated Regression Testing for a Moderating Effect of Changes in Perceived Diversity on the Interrelationship Between Experimental Condition and Changes in the Results of the Public Goods Game

Predictor	Coefficients				
	B	SE	β	t	p
Dummy: Low Diversity Condition	.70	.35	.24	2.02	.047
Dummy: Shared Goal Condition	.97	.34	.33	2.85	.006
Social Desirability	.76	1.20	.07	.64	.53
Δ Perceived Diversity	-.13	.52	-.05	-.25	.80
Low Diversity Condition * Δ Perceived Diversity	1.32	.67	.31	1.96	.053
Shared Goal Condition * Δ Perceived Diversity	.40	.68	.09	.60	.553

Note. Dependent variable was the Δ between the results of the Public Goods Game in the first and in the fourth experimental session. Adjusted $r^2=.175$. $N=91$.

A further interesting effect of including changes in Perceived Diversity as a predictor was that the effect of dummy coded low diversity condition on the results of the Public Goods Game became significant ($\beta=.24$, $t=2.02$, $p=.047$), in contrast to all previous analyses, indicating a suppression effect of Perceived Diversity on the interrelationship between the low diversity condition and changes in the results of the Public Goods Game.

Within a subsequent moderated regression with static Perceived Diversity in the fourth experimental session as an alternative predictor, none of all included predictors were significant (Table 16). This result may imply that Perceived Diversity may be also an influential covariate with regard to the interrelationship between the shared goal condition and the results of the Public Goods Game.

Table 16

Moderated Regression Testing for a Moderating Effect of Static Perceived Diversity on the Interrelationship Between Experimental Condition and Changes in the Results of the Public Goods Game

Predictor	Coefficients				
	B	SE	β	t	P
Dummy: Low Diversity Condition	.40	2.82	.14	.14	.89
Dummy: Shared Goal Condition	-2.75	3.26	-.94	-.84	.40
Social Desirability	.68	1.23	.06	.55	.58
Perceived Diversity Fourth Session	-1.55	1.54	-.22	-1.00	.32
Low Diversity Condition * Perceived Diversity T4	-.23	2.15	-.09	-.11	.92
Shared Goal Condition * Perceived Diversity T4	2.67	2.32	1.27	1.15	.25

Note. Dependent variable was the Δ between the results of the Public Goods Game in the first and in the fourth experimental session. Adjusted $r^2=.061$. $N=91$.

Further analyses on possible moderator, mediator, or suppression effects of changes in Perceived Diversity or static Perceived Diversity on static results of the Public Goods Game in the fourth experimental session revealed no significant effects (Table 17 and Table 18).

Table 17

Moderated Regression Testing for a Moderating Effect of Changes in Perceived Diversity on the Interrelationship Between Experimental Condition and Static Results of the Public Goods Game

Predictor	Coefficients				
	B	SE	β	t	p
Dummy: Low Diversity Condition	.17	.14	.15	1.21	.23
Dummy: Shared Goal Condition	.33	.14	.29	2.36	.021
Social Desirability	-.08	.49	-.02	-.15	.88
Δ Perceived Diversity	-.05	.21	-.05	-.23	.82
Low Diversity Condition * Δ Perceived Diversity	.24	.28	.15	.88	.38
Shared Goal Condition * Δ Perceived Diversity	.22	.28	.13	.81	.42

Note. Dependent variable was the result of the Public Goods Game in the fourth experimental session. Adjusted r^2 = .089. N = 91.

Table 18

Moderated Regression Testing for a Moderating Effect of Static Perceived Diversity on the Interrelationship Between Experimental Condition and Static Results of the Public Goods Game

Predictor	Coefficients				
	B	SE	β	t	P
Dummy: Low Diversity Condition	.53	1.11	.47	.48	.63
Dummy: Shared Goal Condition	-1.36	1.28	-1.19	-1.06	.29
Social Desirability	-.07	.48	-.02	-.14	.89
Perceived Diversity Fourth Session	-.53	.61	-.19	-.88	.38
Low Diversity Condition * Perceived Diversity T4	-.46	.85	-.47	-.54	.59
Shared Goal Condition * Perceived Diversity T4	1.22	.91	1.48	1.34	.19

Note. Dependent variable was the result of the Public Goods Game in the fourth experimental session. Adjusted r^2 = .113. N = 91.

6. Discussion

The first objective of the present study was to broaden the traditional psychological perspective on empathy as the trait of a single person by providing an additional systemic perspective on empathy as the trait of a group. This aim was addressed by applying multiple psychological measures focusing both on individual and systemic empathy, following a multitrait-multimethod matrix (Campbell & Fiske, 1959; Eid, 2000), which was evaluated by partial correlations and confirmatory factor analyses.

The second objective was to investigate the influence of two specific group parameters, diversity and shared goals, on longitudinal changes in individual and systemic empathy. The second aim was addressed within a longitudinal small group design with three experimental conditions. Within one experimental condition, the group parameter “diversity” was manipulated experimentally, whereas the group parameter “shared goal” was manipulated within the second experimental condition. Longitudinal changes in empathy within these groups were contrasted with the longitudinal changes within a control condition by analyses of covariance for repeated measures and multilevel analyses addressing the hierarchical data structure.

The following section provides an interpretation and classification of the results obtained within this investigation. First, the psychometric quality of the applied measures is evaluated, including the measurability of the higher-order factors Individual and Systemic Empathy. The second part addresses the effectivity of the experimental manipulation. The third part provides a detailed discussion of the main findings on longitudinal effects of group parameters on empathy, as well as possible explanations and interpretations of the discovered moderator, mediator, and suppression effects. Fourth, influences of further individual and systemic traits on empathy are evaluated, followed by a part on possible antisocial outcomes of cognitive empathy, which may have been detected within this investigation. Finally, practical implications and limitations of this study are discussed, as well as directions for future investigations on empathy. Based on all mentioned considerations, the dissertation concludes with a final remark on the relevance of this study and future empirical investigations on empathy.

6.1 Psychometric Qualities of the Applied Measures

6.1.1 Objectivity

The objectivity of the procedure within the present investigation was ensured by several means. First, all research assistants and the author of the study participated in an extensive training addressing the correct and standardized application of all measures, the experimental intervention in the second and third experimental session, and technical aspects (e.g., handling of the cameras). Second, the rooms where the experiment was conducted were identical for all participants, as well as the daytime (late afternoon). No external noise was registered during any experimental session. Third, all experimenters acted in accordance with a detailed experimenter manual, including an overview of measures and their correct applications (e.g., how the participants have to be seated to ensure the video supported observation of their nonverbal communication within the DCS; see also Wacker, 2011), standardized oral instructions for all measures, and answers on possible questions by the participants. Further, all questionnaires were preceded by detailed written instructions. The topic of the group discussion was standardized across all experimental conditions. The objectivity of the applied measures and their interpretation is given by standardized evaluation schemes.

In sum, there are no indicators of reduced objectivity within the present study.

6.1.2 Internal Consistency

In general, the internal consistency of the applied measures was satisfactorily high. Out of 59 computed internal consistencies, 44 had an α of .70 or above (range: .49-.95; $M=.77$; $SD=.11$). The internal consistencies of only four subscales were lower than $\alpha=.60$: Perceived Diversity (first session: $\alpha=.51$; fourth session $\alpha=.49$), Diversity Beliefs (fourth session: $\alpha=.49$), and Extrinsic Motivation ($\alpha=.56$). With regard to the subscale Perceived Diversity, the reduced internal consistency is not surprising, as this measure addresses perceived surface-level diversity in the three demographic variables age, gender, and ethnic background. One of these categories may have been perceived more saliently by different participants, based on their own demographic characteristics, personality, and attitudes. For example, a female participant studying Gender Studies may pay more attention to diversity in the variable gender than to diversity in the variable age within her group, whereas the

oldest or the youngest participant within the same group may pay more attention to the age differences than to the gender of the other participants. Based on these considerations, the reduced internal consistency of this measure is rather not critical for the interpretability of the subsequent analyses, especially within Section 5.6, because the main interesting variable is the *mean* perceived surface-level diversity in *all* three demographic variables age, gender, and ethnic background. The concordance between the three single items of the scale may be regarded as still high enough to compute a mean for this subscale, and this mean would even correct for the mentioned bias of individual salience.

The problem of reduced internal consistency is far more severe with regard to the subscale Diversity Beliefs. This result is not in line with the findings on the original scale by Dick et al. (2008) and suggests that the measure is not homogeneous enough to imply a single construct behind it. However, further analyses on this issue would be beyond the scope of this work, therefore the subscale Diversity Beliefs was excluded from follow-up analyses.

The low internal consistency of Extrinsic Motivation in the fourth experimental session is rather surprising, as the consistency was higher in the second ($\alpha=.61$) and in the third ($\alpha=.66$) experimental session, which is rather in line with the reliability of the original scale reported by Cooper and Javatilaka (2006). The most probable explanation for this inconsistently low internal consistency in the fourth experimental session is a random measurement error.

6.1.3 Stability

The stability of the applied measures ranged from zero over moderate up to high stability indices (range: .00-.84; $M=.53$; $SD=.27$; $N=24$). In some cases, e.g. with regard to the subscales Group Identification ($r=.00$) or Group Climate ($r=.04$), stabilities around zero are not problematic and even indicate a higher validity of measurement than higher stabilities, as these scales are not intended to measure systemic traits but rather cognitive and affective states of the group or, simply put, the mood of the group, which is based on highly changeable group dynamics and may represent a relevant covariate in any empirical investigation on group emotions (Carr et al., 2003; Choi et al., 2003). The low stability of the subscale Goal Commitment ($r=.10$) is also not surprising as it may be explained by stronger commitment to systemic goals towards the end of the investigation period, resulting from

naturally occurring developmental processes as the usual increment of group cohesion over time (see standard models of small group development, e.g. Tuckman & Jensen, 1977), and further by shifting perceptions of the central shared goal within a group between the first and the fourth experimental session, as an effect of the experimental manipulation (see Section 5.2.2). The poor internal consistency of the subscale Diversity Beliefs was already discussed, and as it was excluded from further analyses, it may be not of any interest to further discuss its moderate stability ($r=.57$).

However, other reduced stabilities demand further evaluation. The IRI subscales Fantasy ($r=.38$) and Personal Distress ($r=.57$) were constructed to measure stable individual differences in empathy (Davis, 1983), and Davis himself argued that the IRI subscales measure empathy as a trait and supported this statement by reporting considerable stabilities in IRI scores across a re-test interval of even two years (Perspective Taking: $r=.58$; Empathic Concern: $r=.50$; Personal Distress: $r=.59$; Fantasy: $r=.62$; Davis & Franzoi, 1991). The present study revealed a nearly identical (Personal Distress) or much lower (Fantasy) stability across a re-test interval of only three weeks, which may seem surprising at first glance. Random measurement errors within the present study, within the study of Davis and Franzoi (1991), or within both studies may represent simple and possible explanations for this inconsistency. On the other hand, as the stability of the IRI has not yet been investigated across such short re-test intervals as within the present study (at least to the authors knowledge), another reason is also thinkable: Constructs as personal distress or fantasy may not only reflect stable personality traits but may be affected by individual and systemic states to a further extent than the other two facets within the IRI, perspective taking and empathic concern. This is supported by certain empirical investigations, which conceptualize and investigate personal distress as an individual state, being similar to temporary cognitions or worries (e.g., Matthews et al., 2002), and highly dependent on temporary mood states (Curran, Andrykowski, & Studts, 1995), and further by studies on the strong interrelationship between imaginative or fantasizing abilities and mood states (e.g., Baas, De Dreu, & Nijstad, 2008). Of course, these investigations do not exclude the possibility of a stable disposition towards experiencing personal distress or processing imaginative scenarios within specific, interpersonal situations. However, they may explain why the stability of the IRI subscales Personal Distress and Fantasy is substantially lower than the stability of the IRI subscales Perspective Taking and Empathic Concern.

The influence of individual and systemic states may also offer an explanation for the low stability of the DCS group discussion parameters (Empathic Statements: $r=.18$; Empathic Feedback: $r=.53$; Dominance: $r=.55$; Affiliation: $r=.00$) and the Public Goods Game ($r=.47$), which may be more reasonable and appropriate than the influence of random measurement errors. Forgas (1999) conducted two experiments on the interrelationship between mood states and verbal communication strategies and documented strong mood effects on the level of politeness and directness within verbal communication, as well as on further specific communication strategies, which may altogether reflect the measured DCS group discussion parameters to a certain degree. The interrelationship between prosocial orientation, prosocial behavior, and temporary affective states was also documented within numerous investigations (e.g., within the classic study by Isen and Lewin, 1972; but see also recent investigations as Andrade & Ariely, 2009; Capra, 2004; Schulz, Fischbacher, Thöni, & Utikal, 2014).

In summary, there are reasonable explanations for all reduced stability indices within the present study, suggesting that they do not imply random measurement errors or a reduced re-test reliability, but rather reflect actual systemic and individual states within the experimental groups. With regard to the understanding and conceptualization of empathy, these considerations imply that certain facets of the ability to understand and to share the emotions of others cannot be regarded as stable personality traits but rather as transient cognitive and emotional states. This implication is highly remarkable, as it casts doubt on any description or characterization of a personality as empathic, and points to an alternate description of empathy in terms of transient empathic cognitions, emotions, and behaviors.

Further, as the DCS group discussion parameters and the Public Goods Game are systemic empathy measures, the influence of the experimental manipulation, the experimental intervention, and the natural development of systemic traits should not remain unmentioned with regard to explaining the reduced stability of these measures. Traits of a group are based on stable interaction patterns, mental representations of the group members, and interaction scripts (Asendorpf & Wilpers, 1998; see also Section 2.4.1). It is obvious that these determinants of systemic traits could not have existed in the first experimental session, because the group members were unknown to each other. However, as analyses revealed high stability indices for the systemic empathy questionnaire CEEQ-

Group as well as low stability indices for the individual empathy subscales Personal Distress and Fantasy, the significance of this explanation may be limited with regard to the reduced stability indices and rather raises questions on the construct validity of the applied measures, which will be discussed within the following paragraph.

6.1.4 Construct Validity

The first research question within the present study on the concurrent and discriminant validity of systemic and individual empathy measures (see Section 3) was addressed by partial correlations and structural equation modelling in Section 5.4. Out of six confirmatory factor analyses, only one provided support for a two-factor solution with the two latent higher-order factors Individual and Systemic Empathy. Poor model fit indices and anomalies as Heywood Cases (Heywood, 1931) within the remaining five analyses imply that this two-factor solution is invalid with regard to the applied measures within the present study. There are several possible explanations for this result.

The already discussed objectivity and reliability of the applied empathy measures, as well as their high ecological validity resulting from the combination of self-report, image-based, and naturalistic behavioral measures, speak against a strong influence of random or systematic measurement errors within the analyses in Section 5.4. Subsequently, the assumption that the measurement was appropriate directly raises the question whether the constructs systemic and individual empathy actually exist or not. The latter possibility has to be taken into account. However, the presented theoretical considerations on the role of intersubjectivity for the development of empathy (Trevvarthen & Aitken, 2001; see Section 2.3.4), the numerous empirical investigations on dyadic empathy and systemic emotions (see Section 2.4.2), the multiple functions of empathy within a social group (see Section 2.4.3), and the success of psychological interventions which consider the theoretical construct of systemic empathy (e.g., family therapy, see Wilkinson, 1992) support the assumption that systemic empathy exists and that it differs from individual psychological empathy.

Under this assumption, it seems more likely that systemic empathy is too complex to be measured by conventional psychological measures as in the present study. It may be not sufficient to measure empathy on the group level and on the individual level to differentiate between systemic and individual empathy, it could be further necessary to measure

empathy as a dyadic trait (Ickes et al., 1990; Ickes et al., 2000a; Simpson et al., 1995; Stinson & Ickes, 1992) for every possible dyad within a social group, and even the interaction effects between these dyadic effects. This demands a sophisticated measurement method and evaluation, as there are 21 dyads and 210 inter-dyad effects within an exemplary group of only seven persons. Further, triadic effects within a group and the interaction effects between these effects may also be important for a comprehensive measurement of systemic empathy (see also Section 2.4.1). Standard sociometric methods as the Social Network Analysis (SNA; see Scott, 1988; Wasserman, 1994) and even complex statistical approaches as latent space models of social networks (Hoff, Raftery, & Handcock, 2002) seem not sufficient for this purpose, as the mentioned dyadic, inter-dyadic, triadic, and inter-triadic effects have to be evaluated together with individual and systemic levels of empathy within a single matrix. The development of such a new method would go far beyond the scope of this work, but may be addressed by future researchers.

Another even more probable explanation for the reduced construct validity of the applied empathy measures is a strong discrepancy between (1) self-report measures of empathy (IRI, CEEQ-GROUP, MET: Emotional Empathy), (2) performance measures of empathy (MET: Cognitive Empathy), (3) empathic communication (DCS), and (4) associated behaviors (Public Goods Game). Several high correlations between the subscales of the systemic measure CEEQ-GROUP and the individual measure IRI support this conclusion, as well as numerous low or negative correlations between the four enumerated categories of empathy measures (see Section 8.3). This discrepancy between self-report, performance, and behavioral measures of the same construct can be explained by common-method variance, which is not unusual and documented for multiple emotional states and personality traits (see Podsakoff, MacKenzie, Lee, & Podsakoff, 2003, for a review). Brackett, Rivers, Shiffman, Lerner, and Salovey (2006) even documented zero correlations between self-report and performance measures of emotional intelligence, defined as the perception, use, understanding, and management of emotions, and therefore being close to the operationalization of empathy in the present study. However, to calculate and correct for the common method variance it would have been necessary to include additional measures within the present study, e.g. a systemic image-based measure of empathy, which would require images with emotional facial expressions of all participants, and the measurement of individual communication parameters and individual behavioral patterns within economic

games, which would in turn require multiple measurements within different relational constellations. Unfortunately, it was not possible to develop and to include further and even more effortful measures within the present investigation.

Nonetheless, even if the construct validity of the applied measures and therefore the interpretability of the main analyses may be limited, it might be interesting and fruitful to discuss the main analyses separately with regard to the distinct categories of self-report measures, performance measures, empathic communication, and supportive behavior. However, before proceeding with discussing the longitudinal effects of experimentally manipulated group parameters on empathy, the effectivity of the experimental manipulation should be evaluated.

6.2 Effectivity of the Experimental Manipulation

The experimental manipulation within the present investigation had several objectives. First, it aimed to induce significantly lower levels of perceived surface-level diversity within the low diversity condition than within the control condition and within the shared goal condition. Second, it aimed to implement a measurable cognitive representation of a shared goal within the shared goal condition, but not within the control condition and within the low diversity condition. Additionally, participants within the shared goal condition were meant to be committed to their shared goal, to be intrinsically rather than extrinsically motivated to pursue it, and to be oriented towards cooperation rather than towards competition in pursuing it. The manipulation checks in Section 5.2 suggest that every objective of the experimental manipulation was achieved. However, the success of this specific experimental manipulation within the present study is not a matter of course and remarkable for several reasons.

As already stated, the effects of objective indicators of surface-level diversity within a social group are highly dependent on their perception and interpretation by the group members (e.g., Stürmer et al., 2006; see Section 2.4.4.1). Several empirical investigations suggest that this perception and interpretation of surface-level diversity may be not only dependent on variables as diversity beliefs (van Dick et al., 2008; see Section 4.3.1.6), but also on various other states and traits, which cannot be manipulated or controlled for easily. Most notably, Meyer, Shemla, and Schermuly (2011) demonstrated that the perception of

diversity, which affects team performance and information elaboration, is moderated by the subjectively perceived salience of a social category. For example, it is possible that the *perceived gender diversity* within an exemplary social group is low, even if the *objective gender diversity* is high, because the mental representations of the members within this social group do not incorporate gender as a salient and relevant category. Another study by Blanz (1999) revealed that the salience of social categories as gender is not only dependent on this kind of issue relevance and situational accessibility, but also on the meta-contrast ratio between intra- and intercategory differences: Within an exemplary group with high levels of objective gender diversity, the perceived gender diversity will be the higher, the smaller the differences *within* both genders are, e.g. with regard to other characteristics as personality or age. Such results imply that it was risky to manipulate the objective surface-level diversity in age, gender, and ethnic background within the present study without additionally triggering the salience of these social categories (for example by addressing general gender topics within the group discussion) and without manipulating the meta-contrast ratio, which would have been extremely difficult, if not impossible for all three social categories. However, as the experimental manipulation of perceived diversity within this investigation was successful without these additional procedures, it demands further evaluation. The first reason for this success may be a critical difference between the present investigation and the study of Meyer et al. (2001): The authors used not only typical indicators of surface-level diversity as gender to manipulate diversity faultlines, but also indicators of deep-level diversity as personality assessments. Blanz (1999) describes surface-level indicators as social categories with “*high chronic accessibility*” (p.1), and a categorization of other humans by gender, age, and ethnicity occurs within milliseconds and immediately activates associated stereotypes (Fiske & Russell, 2010). Thus, these three social categories could be not only chronically accessible, but by tendency also chronically salient, in contrast to other social categories as personality traits. The second reason for the success of the experimental manipulation may be specific sample characteristics, which advantageously influenced the meta-contrast ratio. For example, as all participants were students with higher levels of education, the differences *within* the manipulated categories gender, age, and ethnicity were rather small. These factors have to be considered within psychological interventions (e.g., group counseling; see Corey, 2012) and future empirical investigations attempting to manipulate the perceived diversity within a social group.

The successful implementation of a cognitive representation of a shared goal, as well as the induction of commitment, intrinsic motivation, and a cooperative orientation are also noteworthy. To the knowledge of the author, no psychological or sociological study on goals achieved a successful experimental manipulation of this specific but for investigative purposes highly desirable combination of states. Moreover, the general tendency of study participants to minimize their cognitive and behavioral effort within experiments is frequently discussed as one of the most critical challenges for experimental psychology, because it heavily affects the validity of all results (e.g., Krosnick, Narayan, & Smith, 1996; Krosnick, 1999; Oppenheimer, Meyvis, & Davidenko, 2009). The fact that the present investigation was not affected by this specific experimental bias may be explained by several interventions and circumstances. First, the instructions on shared and individual goals were extensive and frequently repeated (see Section 4.2.1 for details). This procedure increased the probability that every participant understood the instructions, cognitively processed them, and finally generated the mental representation of a shared or individual goal. Second, the relevance of and commitment to the shared goal within the corresponding experimental condition was increased by initially announcing an evaluation of the created artwork at the end of the investigation period and the possibility of exhibiting it afterwards. Third, the specific goal of creating a creative collage as a moderately difficult task may have increased goal commitment to a further extent than easy and difficult tasks (Wright et al., 1996; see Section 4.2.1). And fourth, the possibility of exhibiting the artwork probably elicited the perception of an intergroup competitive setting within the shared goal condition, which usually triggers an intragroup cooperative orientation, and an intragroup competitive setting within the low diversity and within the control condition (Mulvey & Ribbens, 1999; van Mierlo & Kleingeld, 2010). These details are not only relevant with regard to future investigations on personal goals, but also with regard to various psychological intervention methods that try to implement shared goals within a social group, e.g. intergroup interventions based on the Intergroup Contact Hypothesis (Pettigrew & Tropp, 2005) or Jigsaw Techniques (Aronson, 1978).

In sum, several important implications for psychological science and practice are associated with the noteworthy success of the experimental manipulation within the present study.

6.3 Longitudinal Effects of Group Parameters on Empathy

The following section provides an interpretation and explanation of the longitudinal effects of group parameters on empathy, which were analyzed in Section 5.5, and the associated moderator, mediator, and suppression effects, which were revealed in Section 5.6.

6.3.1 The Impact of Diversity on Empathy

With regard to the two self-report measures IRI and CEEQ-GROUP, as well as to the Multifaceted Empathy Test and the DCS group discussion parameters, the analyses of covariance for repeated measures and multilevel models in Section 5.5 did not reveal any significant differences between the low diversity and the control condition, except for the CEEQ-GROUP subscale Mirroring. However, as this significant effect is completely opposite to the hypothesized differences between the experimental groups and was only found within the ANCOVA for repeated measures and not confirmed within the subsequent multilevel analysis, it is most probably caused by random group-level variance and will not be discussed further.

Significant differences between the low diversity and the control condition in scores of the individual empathy measures IRI and MET would have been surprising and not in line with the hypotheses of the present study (see Section 3), as individual empathy was not addressed by the elaborate intervention program of this investigation (see Section 4.2.2 and Section 4.2.3). However, the finding that changes in systemic empathy scores did also not significantly differ between the low diversity and the control condition, despite of a successful manipulation of perceived surface-level diversity, is contrary to the first hypothesis of this investigation.

Considering the limited construct validity (see Section 6.1.4), further important limitations of this study (most importantly the short investigation period; see Section 6.7 for a detailed discussion), and the revealed suppression effect of perceived diversity on the interrelationship between surface-level diversity and prosocial behavior within the Public Goods Game (see Section 5.6), it is necessary to interpret this finding with extreme caution. Nonetheless, after adopting this cautious attitude, it is inevitable to discuss the possibility that the results on the systemic measures CEEQ-GROUP and the DCS group discussion parameters reflect the truth. If this assumption is correct, it would be not only necessary to

reject the first hypothesis, but also to revise several theories and empirical findings on the interrelationship between empathy and similarity (e.g., Cialdini et al., 1997; Stürmer et al., 2006; see Section 2.4.4.1). The first step towards such a revision is a critical evaluation of the main differences between previous investigations on this interrelationship and the present study.

First, the experimental design of previous investigations on the interrelationship between empathy and similarity was not as naturalistic as in the present study. These studies, even if motivating the formulation of the first hypothesis, did not measure empathy and prosocial behavior directly within face-to-face interactions but by imaginative, fictitious scenarios (Cialdini et al., 1997) or by reactions on preprogrammed actions and traits of fictitious participants via e-mail (Stürmer et al., 2006). However, a substantial body of evidence suggests that the imagination of empathy with fictitious persons may strongly differ from empathy within an actual social context, and that helping intentions may strongly differ from actual helping behavior. A specific difference between imaginations of social interactions and the actual behavior within such situations may explain why previous studies found an interrelationship between empathy and similarity within imaginative scenarios, which is not existent in observed social behavior: The imagination of social situations which are familiar demands less cognitive elaboration and occurs more often than the imagination of social situations which are not familiar (Sharman, Garry, & Hunt, 2005). As the personality traits and physiological parameters of most interaction partners in our daily routine are positively correlated to ours, which can be partially explained by genetic similarity in case of family members, and by active as well as passive partner selection effects in case of romantic relationship partners and friends (see Asendorpf & Banse, 2000), it is obvious that most persons understood and shared the emotions of similar persons far more often than the emotions of dissimilar persons. Therefore, the imagination of empathy with a similar person demands less cognitive elaboration and is easier than the imagination of empathy with a dissimilar person. As this argumentation is limited to imaginations and not transferable to observable behavior, results on the interrelationship between empathy and similarity are highly questionable, if they are based on imaginative scenarios only. This question of ecological validity must also be considered with regard to the measurement of supportive behavior: Several studies documented a strong discrepancy between self-reported behavior in imaginative social situations and observed behavior in actual social

situations (see the classical study by LaPiere, 1934, and the meta-analysis by Wicker, 1969), and a recent meta-analysis revealed an additional discrepancy between subjective norms, intentions, and actual behavior (Armitage and Conner, 2001). In other words, subjective and socially desirable norms cannot be equated with planned behavior, and planned behavior cannot be equated with actual behavior.

This differentiation points to the second critical difference between the present study and previous investigations on the interrelationship between empathy and similarity: The findings of these previous investigations rely on self-report measures only. However, as empathy and prosocial behavior are socially desirable, it is highly important to consider and statistically control for the effect of social desirability on self-reported empathy and prosocial behavior (e.g., Singer & Lamm, 2009; Dziobek et al., 2008).

Third, no previous investigation on the interrelationship between empathy and similarity measured empathy as an emergent, systemic trait. The importance of this additional perspective on empathy, as well as the differences and interaction effects between individual, dyadic, and systemic empathy were addressed in Section 2.4. As the present investigation further differentiates between cross-sectional (see Section 5.1) and longitudinal effects of relationship parameters on individual and systemic empathy, it could be associated with a higher level of incremental validity than previous studies.

Fourth, the experimental design of Cialdini et al. (1997) entailed additional variables possibly interfering with empathy and perceived similarity, as the degree of kinship. The authors measured imagined empathic concern and prosocial intentions towards strangers, acquaintances, friends, and family members. The closer the imagined relationship was, the higher was the degree of imagined empathic concern and prosocial intentions, and this association was mediated by the degree of perceived similarity, which was measured by a oneness index. The authors concluded that it is easier to understand and to share the emotions of similar persons than to understand and to share the emotions of dissimilar persons. However, as relationship closeness, partially operationalized by the degree of kinship, and the degree of perceived similarity were highly correlated within this study ($r=.66$), they may not represent two independent variables. As a consequence, the authors may have measured kinship orientation instead of perceived similarity. The strong prosocial orientation towards genetically related persons is a well-known and already documented evolutionary principle (e.g., Neyer & Lang, 2003), but it is possible that the interrelationship

between prosocial orientation and similarity is weaker or even not existent with regard to genetically unrelated persons: Even if perceived similarity may serve as an indicator of genetic similarity, this proximate mechanism may be limited to genetically related persons. Although there is no empirical evidence on this thesis, it seems important to differentiate between empathy towards genetically related and genetically unrelated persons.

However, as already mentioned, the consideration of these main differences between previous investigations and the present study is only the first step towards a revision of conventional positions on the interrelationship between empathy and similarity, as it primarily addresses methodical issues. The second step is a thorough content-related evaluation of possible reasons for why there could be no association between empathy and perceived surface-level similarity.

The phylogenetic perspective on empathy (see Section 2.3.1; Plutchik, 1987) may provide an explanation for finding neither cross-sectional nor longitudinal effects of perceived similarity on the level of empathy within the present investigation. If we adopt an evolutionary view on empathy, the ability to understand and to share the emotions of similar as well as of dissimilar persons may be associated with survival values: Every natural social group is characterized by a certain degree of dissimilarity between the group members, especially with regard to the surface-level indicators age and gender. Consequently, a social group where group members only recognize and share the emotions of similar group members will be not as efficient in gathering food, hunting, and defending against predators or natural hazards as a social group where group members ubiquitously recognize and share the emotions of others. Even with regard to out-group members it may be beneficial and even vital to recognize the emotions of dissimilar persons, as it is a precondition for the anticipation, response to and prevention of threatening or violent behaviors of out-group members. This principle may be valid for the surface-level indicators age and gender only, and not for ethnic background or indicators of deep-level diversity. Unfortunately, this kind of differentiation and the additional manipulation of deep-level diversity would go far beyond the scope of this work.

In this context, the mentioned differentiation between empathy towards in-group and towards out-group members may be more important than a differentiation between empathy towards similar and dissimilar persons. The discussed function of anticipating

violent behavior is only limited to cognitive empathy, emotional empathy towards in-group members is probably associated with higher survival values than emotional empathy towards out-group members, and empathic responses towards in-group members occur more frequently and are therefore more relevant fitness indicators than empathic responses towards out-group members. This perspective would be in line with the findings of Thibault et al. (2006), who documented lower levels of empathy towards out-group members than towards in-group members after controlling for similarity and expression styles. To investigate this difference within the present investigative design, it would have been necessary to measure individual, systemic, and inter-systemic empathy. The measurement of intergroup emotions and its complexity was reviewed in Section 2.4.2 and was also not possible in the present study.

On the other hand, if there is actually no interrelationship between empathy and perceived similarity, how to explain the numerous studies documenting that social groups with high levels of surface-level diversity have lower levels of cohesiveness, social integration, perform worse in various tasks, and have higher levels of interpersonal conflict than social groups with low levels of surface-level diversity (Harrison et al., 1998; Levine & Moreland, 1990; O'Reilly et al., 1989; van Knippenberg et al., 2004)? First, as empathy was not measured within these investigations, it is possible that other individual or systemic variables determine the interrelationship between the mentioned group parameters and surface-level diversity, most probably indicators of deep-level diversity as the level of cognitive heterogeneity. Second, Harrison et al. (1998) revealed that the effect of surface-level diversity on a group's cohesion was significantly weaker after a period of six weeks, while the level of deep-level diversity was significantly stronger after this period, probably because regular meaningful interactions reduce the salience of surface-level diversity indicators and enhance the salience of deep-level diversity indicators as personality traits. This process may have influenced the longitudinal effects of surface-level diversity on empathy within the present study. Third, a meta-analysis by Webber & Donahue (2001) correcting for artifactual variance did not find any interrelationship between neither surface-level diversity nor deep-level diversity and a group's cohesion or performance. This finding is in line with the results of the present investigation and the phylogenetic perspective on empathy.

In sum, even if the first hypothesis has to be rejected, there are several reasonable methodological and theoretical arguments for why there could be no association between empathy and perceived similarity at all, and there is some empirical evidence (Webber & Donahue, 2001) indirectly supporting this assumption.

To conclude this section, the suppression effect of changes in the subscale Perceived Diversity on the interrelationship between the low diversity condition and changes in the results of the Public Goods Game (see Section 5.6) should be discussed.

Even if the analyses in Section 5.5.3 revealed no significant effect of diversity within a social group on the changes in the results of the Public Goods Game between the first and the fourth experimental session, the effect of diversity became significant after entering changes in the subscale Perceived Diversity into a moderated regression equation. In statistical terms, such a classical suppression effect can be explained as follows: If two predictors as the low diversity condition and the subscale Perceived Diversity are substantially interrelated (see Section 5.2 and 6.2), irrelevant variance of the main predictor (low diversity condition) is partialled out after including the additional predictor, which is further indicated by its negative regression weight after entering it into the regression equation (see Section 5.6, Table 15; Maasen & Bakker, 2001). The suppressant (changes in perceived diversity) is associated with the predictor (low diversity condition), but not with the dependent variable (changes in the results of the Public Goods Game; $\beta = -.05$; $t = -.25$; $p = .80$). However, a suppression effect occurs only if the main and the additional predictor have commonalities, which are irrelevant for the association between the main predictor and the dependent variable. For example, the negative association between self-esteem and antisocial behavior increases after entering narcissism into the regression, because narcissistic, maladaptive, and therefore in this case irrelevant forms of self-esteem are partialled out by this procedure (Paulhus, Robins, Trzesniewski, & Tracy, 2004). But what commonalities between the dummy-coded low diversity condition and changes in perceived diversity could be irrelevant and distorting with regard to the effect of diversity on changes in prosocial orientation within the Public Goods Game? Most probably, these irrelevant commonalities were increasingly salient deep-level diversity parameters, which were less relevant in the first than in the fourth experimental session: The manipulation check (see Section 5.2 and 6.2) revealed significantly lower levels of perceived diversity within the low

diversity condition than in the other experimental conditions, both in the first and in the fourth experimental session. Finding such a difference within the first experimental session can be explained by the high chronic accessibility of surface-level indicators as age or gender (Blanz, 1999; Fiske & Russell, 2010). If a person meets other persons of the same age, gender, and ethnic background, it is reasonable that this person perceives them as similar at first glance. However, each interaction between the members of such a social group offers the possibility to discover dissimilarities in deep-level diversity parameters as personality traits (Harrison et al., 1998). After a period of three weeks involving regular interactions between the group members it is highly probable that some groups within the low diversity condition discovered more deep-level dissimilarities than other groups, even if all groups within the low diversity condition had lower levels of perceived diversity than groups within the other experimental conditions, both in the first and in the fourth experimental session. Mean differences in the subscale Perceived Diversity between the first and the fourth experimental session may reflect this process of discovering deep-level dissimilarities within groups in the low diversity condition. Thus, including these mean differences as an additional predictor could partial out this confounding variance and strengthen the predictive value of the experimental condition on changes in the results of the Public Goods Game. Unfortunately, it is impossible to verify this assumption without measuring indicators of deep-level diversity and their perception.

The presented interpretation of the suppression effect implies that perceived surface-level similarity is associated with prosocial orientation within a social group, if changeable effects of perceived deep-level similarity are partialled out. This implication is in line with the first hypothesis of this investigation, but not with the results on the other systemic empathy measures as well as the presented arguments against an association between empathy and perceived similarity. Most probably, this ambiguity can be finally resolved by additional longitudinal investigations including further systemic and individual empathy measures (see Section 6.1.4) and measuring indicators of deep-level diversity and their perception.

6.3.2 The Impact of Shared Goals on Empathy

The results on the interrelationship between shared goals and empathy follow a pattern, which is very similar to the reported findings on the interrelationship between

diversity and empathy. With regard to the individual empathy measures IRI and MET, there were no significant differences between the shared goal condition and the control condition, neither within the analyses of covariance for repeated measures nor within the multilevel analyses (Section 5.5), which is in line with the second hypothesis of the present investigation. On the other hand, there were also no significant differences with regard to the systemic empathy measure CEEQ-GROUP and the DCS group discussion parameters, despite a successful experimental manipulation of goal representation, goal commitment, intrinsic motivation, and cooperative orientation (see Section 5.2), which is not in line with the second hypothesis. However, there was a significant difference between the shared goal condition and the control condition with regard to the results of the Public Goods Game: The increase in the amount of money that was paid into the common cash box occurring between the first and the fourth experimental session was significantly higher within the shared goal condition than in the control condition (see Section 5.5.3). This effect is in line with the second hypothesis, and further mediated by group climate and moderated by the level of group identification (see Section 5.6). In sum and under consideration of the latter result, the level of confirmation for the second hypothesis can be regarded as slightly higher than for the first hypothesis. Similarly as in the previous section, and independently from the limitations of the present investigation (see Section 6.7), it is important to compare these results with previous findings on the interrelationship between the cognitive representation of a shared goal and empathy-related states and traits within social groups (see Section 2.4.4.2), which motivated the formulation of the second hypothesis.

The findings by Wegge and Hasslam (2005), who reported that the cognitive representation of a shared goal within a social group is associated with higher levels of group identification, and by Aubé and Rousseau (2005), who documented that the cognitive representation of a shared goal within a social group elicits supportive behaviors towards other group members, were both clearly confirmed within the present study. Further, the analyses in Section 5.6 documented an interdependence between both effects: The association between the perception of shared goals and the emergence of supportive behaviors within social groups is moderated by the level of group identification. The revealed association between group identification and supportive behaviors is also in line with the empirical study by Simon et al. (2000; see Section 4.3.1.5). Without considering the

results on the systemic empathy measures CEEQ-GROUP and the DCS group discussion parameters, it may seem reasonable to explain this moderation effect by extending the model of Cialdini et al. (1997): If a person realizes that he or she shares a goal with another person, a feeling of oneness between both persons may emerge – the other person may be perceived as more similar and it may be easier to identify with this person than without the perception of a shared goal. Subsequently, higher levels of perceived similarity and identification with the other person may foster empathy, and empathy may in turn elicit prosocial behavior (see Section 2.1.3).

However, the arguments against a substantial interrelationship between perceived similarity and empathy (see Section 6.3.1) point to an alternative explanation. A hint towards such an alternative model is the finding that the interrelationship between shared goals and supportive behaviors within groups is not only moderated by the level of group identification but also mediated by group climate (see Section 5.6). This result on the role of group climate does not only indirectly confirm the results of Klein et al. (2001; see Section 2.4.4.2), who reported that the perception of a shared goal within a social group is associated with positive affective reactions towards other group members, and of Haas et al. (1992), who revealed that groups with perceived shared goals were more satisfied with their communication than groups without perceived shared goals. The mediator effect most importantly indicates that prosocial and supportive behaviors may result from general positive affect and that this process can be completely independent from genuine empathy.

On the one hand, this interpretation challenges the empathy-altruism hypothesis (Batson et al., 1981, 1988, and 1991) and the classical functional perspective on empathy as a main trigger of prosocial behavior (see Section 2.1.3). On the other hand, it would be fully in line with the results on the systemic empathy measures CEEQ-GROUP and the DCS group parameters, as well as with the arguments against an interrelationship between perceived similarity and empathy in the previous section. Furthermore, there are two additional arguments supporting this alternative perspective on the emergence of supportive behaviors within social groups.

The first argument addresses the methodology of the central investigations, which motivated the formulation of the empathy-altruism hypothesis: Batson et al. (1981; 1988; 1991) did not operationalize and measure empathy (see Section 2.1.3), their thesis was

based on qualitative interpretations and attributions only. Moreover, they did not control for the role of general positive affect within their investigations, which may have been an alternative variable explaining altruistic behavior.

The second argument is the already discussed low stability of prosocial orientation within the Public Goods Game ($r=.47$; see Section 6.1.3). This result already indicated that prosocial orientation within groups is highly dependent on states, and the mediator analysis (see Section 5.6) finally revealed on which state exactly: a positive group climate, or, in other words, high levels of general positive affect within a social group. This explanatory model refers to a social system but is substantially supported by individual psychological research on the role of positive affect for explaining human prosocial behavior (Andrade & Ariely, 2009; Capra, 2004; Isen & Lewin, 1972; Schulz et al., 2014). The present investigation suggests that the cognitive representation of a shared goal can elicit systemic positive affect within a social group.

As a matter of course, explaining prosocial behavior within the groups of the present study by general positive affect and not by empathy does not imply that there is no association between empathy and prosocial behavior, which was documented within numerous behavioral, developmental, evolutionary, and neuroscientific studies (Baaren et al., 2004; Decety & Jackson, 2004; de Vignemont & Singer, 2006; Eisenberg & Fabes, 1990; Preston & de Waal, 2002; see also Section 2.1.3). However, similarly as Batson et al. (1981; 1988; 1991), these studies did not control for the level of general positive affect. The present study demonstrates that temporary affective states have to be taken into account with regard to explaining prosocial behavior and that the link between empathy and prosocial behavior may be not as strong as it is classically postulated.

Further, the presented interpretation does not imply that there is no association between empathy and general positive affect, which was indicated by a recent biological study on mice and men: Martin et al. (2015) demonstrated that a social encounter with strangers is associated with higher levels of stress than an encounter with familiar subjects and that this form of social stress usually inhibits emotional contagion of pain in stranger dyads. Social stress is usually regarded as a negative affective state (e.g., Curran et al., 1995; Matthews et al., 2002). However, specific social situations, e.g. a shared gaming experience,

were being shown to reduce social stress and to enable emotional contagion in stranger dyads, which is comparable to emotional contagion in familiar dyads (Martin et al., 2015). This effect was found both in human and in mouse dyads.

It is highly probable that this mechanism is not only limited to dyads: When persons initiate interactions with strangers within a social system, group parameters as the cognitive representation of a shared goal and cooperative orientation may elicit general positive affect, and this affective state may in turn enable emotional contagion between all group members, as if the group members were familiar with each other. This mechanism may finally enable genuine empathy, as emotional contagion is an important precursor of empathy (Hoffmann, 2000; Singer & Lamm, 2009; see Section 2.2.2).

Nonetheless, the present investigation demonstrates that this path from perceiving shared goals to general positive affect to emotional contagion to genuine empathy is a protracted process, which rather cannot be completed within three weeks after group formation. The experimental groups within the shared goal condition completed the path from perceiving shared goals to general positive affect and the study by Martin et al. (2015) suggests that they also completed the path from general positive affect to emotional contagion, even if the latter construct was not directly measured within the present investigation. However, the results on the systemic empathy measures CEEQ-GROUP and DCS group discussion parameters indicate that these groups did not take the final step from emotional contagion to genuine empathy, involving the activation of higher-order processes as empathic concern or abstract reasoning on the emotional states of other group members, as well as the shift from self-focus within emotional contagion to other-focus within genuine empathy (see Section 2.1 and 2.2.2). This initial inhibition of empathy with strangers seems to be strong, as it was observable despite of the elaborate intervention program (see Section 4.2), which was designed to accelerate group formation processes in all experimental groups. In the following, it may be interesting to evaluate possible reasons for this initial inhibition of empathy, which seems to occur despite of several important precursors as cooperative orientation, general positive affect, and emotional contagion.

Once again, the probably most reasonable argument supporting the revealed inhibition addresses the phylogenetic perspective on empathy (see Section 2.3.1). Most likely, simple precursors of empathy as emotional contagion and the basic understanding of

facial expressions already realize the crucial survival values of empathy within social groups, e.g. the enablement of coordinated activities, quick responses to predators or natural hazards (Plutchik, 1987), and adequate reactions to an infant's emotional states. The additional, other-oriented higher-order processes of genuine empathy, e.g. empathic concern and abstract reasoning on the other person's emotional states, most likely entail a far-reaching relational investment, which may, in extreme cases, result in transferring survival values from the empathizing person to the target with whom this person empathizes. Within long-term relationships, characterized by high levels of trust, this strong form of investment may be in most cases highly valuable for both parties, as it results in higher levels of commitment and functionality. Within short-term relationships or stranger dyads, the same investment must be regarded as extremely risky: A person empathizing with a stranger delivers him- or herself up to the hands and intentions of an unknown person with unknown motives. The results of the present study indicate that this kind of risky investment was usually not rewarded within human phylogenesis, because, most probably and in most cases, it resulted in the described shift of survival values. Moreover, it is possible that this shift can be enforced by the unknown person's deceptive strategies aiming to take advantage of the empathizing person. These deceptive strategies, e.g. faking pain expressions or distress to gain extensive social support, may in turn be associated with survival values. This association between cognitive empathy and antisocial exploitation will be further discussed in Section 6.5, as it is indirectly confirmed by some specific, even though exploratory analyses of the present study.

This phylogenetic argumentation is consistent with a recent investigation by Liebal, Vaish, Haun, & Tomasello (2014) on great apes, who did not observe empathic reactions to other apes within stranger dyads, while several previous observational studies identified such behaviors in great apes (e.g., de Waal, 2008; see Section 2.3.1).

However, the described initial inhibition of empathy in short-term relationships does not imply that prosocial and altruistic behaviors are generally inhibited, too. The participants within the shared goal condition of the present investigation did not have any individual advantage of sharing their money with the other group members by giving it into the common cash box within the Public Goods Game. However, the cognitive representation of a shared goal elicited general positive affect, and this affect elicited prosocial and unselfish

behavior, even without experiencing genuine empathy. It is possible that temporary affective states serve as a heuristic indicator for how risky relational investments within short-term relationships are, whereas empathy persistently fosters such relational investments in long-term relationships.

Summing up, the main conclusion of this section is that general positive affect is an important predictor of prosocial and altruistic behavior within short-term relationships, and that empathy is an increasingly important predictor of prosocial and altruistic behavior within long-term relationships. Unfortunately, it would have been necessary to study the experimental groups of the present investigation over a longer period of time than three weeks to verify this conclusion. The aspect of the short investigation period will be further evaluated in Section 6.7.

6.4 Correlates of Empathy

The previous section addressed the confirmatory analyses on the two hypotheses of this dissertation. Even if other analyses (especially the correlations reported in Section 5.3) have exploratory character and have to be interpreted with more caution than confirmatory analyses, it would be a great loss to ignore them completely, as they may be relevant for several future investigations on individual or systemic empathy. The following section will therefore evaluate meaningful associations between individual and systemic empathy on the one hand and several individual and systemic states and traits on the other hand, which were revealed in the present study. These findings refer to the total sample and are independent from the three experimental conditions.

6.4.1 Individual States and Traits

The interrelationship between social desirability and the empathy measures of the present investigation (e.g., the IRI subscale Perspective Taking, the CEEQ-group subscale Perspective Taking, and the DCS group discussion parameters; see Section 5.3.1) was already addressed within Section 6.3.1 (Dziobek et al., 2008; Singer & Lamm, 2009) and has been previously reported by several researchers (e.g., by Watson & Morris, 1991, who reported a slightly higher correlation coefficient of $r=.38$ between the IRI subscale Perspective Taking

and a measure of social desirability). The results of the present investigation confirm that the ability to understand and to share the emotions of others is a socially desirable trait and further emphasize the importance of statistically controlling for socially desirable responses within any analysis on self-report measures of empathy.

The associations between the Big Five personality traits and the empathy measures of the present investigations (see Section 5.3.2) provide much more interesting insights.

To the knowledge of the author, the revealed substantial correlation between Neuroticism and the IRI subscale Personal Distress was not reported within previous investigations, and a possible association between emotional instability and empathy was not addressed within any theoretical model of the ability to understand and to share the emotions of others. This lack of evidence leads to the question, if it is appropriate to conceptualize personal distress as an integral facet of empathy, as it was stated in Section 2.1.2 (see also Davis, 1980; 1983). In fact, some researchers differentiated between personal distress as a self-oriented reaction and empathy as an other-oriented reaction (e.g., Eisenberg et al., 1989). However, it seems highly reasonable that even strictly other-oriented empathy is associated with substantial costs for the empathizing person, because sharing and being aware of another person's negative emotional states (e.g., severe pain or despair) pushes both mind and body into a state of alert. These substantial costs may be most appropriately operationalized by the construct personal distress, and general emotional instability may be the outcome, the cause, or both outcome and cause of personal distress. This interrelationship further supports the previous interpretation and conceptualization of empathy as a far-reaching relational investment (see Section 6.3.2) and suggests that the ability to understand and to share the emotions of others is not only associated with beneficial outcomes, especially with regard to the empathizing person. This perspective seems to be rather underrepresented within psychological and sociological research on empathy, despite its practical relevance for several professions, e.g. therapists and physicians. In sum, the reported substantial correlation signifies the role of personal distress as an integral component of empathy.

The reported associations between Extraversion and empathy are interesting, as there are both positive correlations, e.g. between Extraversion and the IRI subscale Fantasy and the CEEQ-group subscale Mirroring, and negative correlations, e.g. between

Extraversion and the MET subscale Cognitive Empathy. Most probably, this discrepancy can be explained by the difference between self-report and naturalistic performance measures of empathy: The self-image of highly extraverted persons seems to be characterized by high levels of intuitive emotional mimicry and imaginative perspective taking, but when it comes to recognize and to understand facial expressions in a naturalistic setting (see Section 4.3.2) they perform worse than less extraverted persons. This discrepancy between self-image and observable behavior may be caused by an overrepresentation of histrionic personality traits within highly extraverted persons, which was already documented in several studies (see Saulsman & Page, 2004, for a meta-analytic review). This result further exemplifies the importance of applying naturalistic performance measures of empathy in addition to self-report measures and demonstrates the limitations of previous studies on the interrelationship between the Big Five personality traits and empathy, e.g. of the study by van der Zee et al. (2002), who reported a positive association between Extraversion and empathy based on questionnaire measures. Finally, this finding further confirms the method-based explanation for the reduced construct validity of individual and systemic empathy measures in Section 6.1.4.

Further, analyses revealed high positive correlations between the Big Five personality trait Openness and self-report measures of cognitive empathy, e.g. the IRI subscales Perspective Taking and Fantasy. In contrast to the previously reported correlations between Extraversion and the MET subscale Cognitive Empathy, Openness was positively associated with this subscale, and further with the DCS group discussion parameter Empathic Feedback. Thus, the reported correlations indicate congruence between the self-image of persons with high levels of Openness and the observable behavior of these persons. The association between Openness and both self-reported and observable empathy can be explained by two arguments. First, the ability to recognize, to understand, and to reflect upon the emotions of other persons is not only a relational investment (see Section 6.3.2) but also a question of individual skill (Constanzo, 1992; Dziobek et al., 2008) and therefore most probably dependent on general intelligence. It seems highly reasonable that persons with high levels of general intelligence have better abilities to decode and interpret the emotions of others than persons with low levels of general intelligence. This interrelationship is in turn the theoretical background of the construct Emotional Intelligence (e.g., Brackett et al., 2006), operationalized as the perception, use, understanding, and management of emotions. As the

Big Five personality trait Openness is substantially associated with general intelligence (e.g., McCrae, 1994; Moutafi, Furnham, & Crump, 2006), it is also reasonable that Openness and both self-reported and observable empathy are interrelated. The second argument refers to commonalities between empathy and openness, which are independent from general intelligence. On the one hand, several definitions of empathy-related constructs are explicitly based on openness to another person's emotional states, e.g. the definition of compassion (Gilbert, 2005; see Section 2.2.3). If two persons are not open-minded towards each other, it is hardly possible that they mutually understand and share their emotions. This perspective may be further helpful in explaining why facets of both cognitive and emotional empathy (e.g., the MET subscale Emotional Empathy) are associated with Openness. On the other hand, the Big Five personality trait Openness entails facets as curiosity, sensitivity, and fantasy (Costa & McCrae, 1992), being similar to the multifaceted structure of empathy (see Section 2.1.2). Moreover, being generally open to new experiences most probably entails being open to the emotional states of other persons, as the adoption of an alternate emotional pattern and perspective can be regarded as a new experience. In sum, these arguments point to a possible overlap between the three constructs empathy, openness, and general intelligence, which can be derived from Figure 27. However, as all analyses in Section 5.3 have exploratory character, they demand further confirmatory investigations.

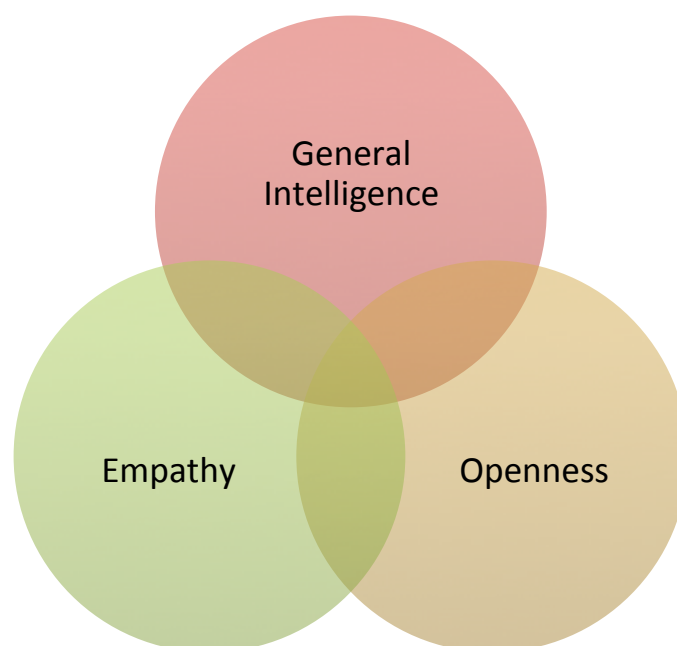


Figure 27. Hypothesized Overlap Between Empathy, Openness, and General Intelligence.

The Big Five personality trait Conscientiousness was associated with two empathy measures in the first experimental session (the IRI subscale Perspective Taking and the CEEQ-GROUP subscale Perspective Taking), and with six empathy measures in the fourth experimental session, including self-report measures, naturalistic performance measures, and behavioral measures of cognitive as well as of emotional empathy (the IRI subscale Perspective Taking, the MET subscale Emotional Empathy, the CEEQ-GROUP subscale Mental State Perception, the CEEQ-GROUP subscale Empathic Concern, the CEEQ-GROUP subscale Mirroring, and prosocial orientation within the Public Goods Game). These results indicate that throughout the initial period of group formation, the individual level of Conscientiousness becomes an increasingly important predictor of both self-reported and observable empathy. From a methodological point of view, this implication demonstrates the limits of previous cross-sectional studies on the interrelationship between empathy and Conscientiousness, revealing negligible (Barrio, Aluja, & Garcia, 2004) or even zero correlations between empathy and Conscientiousness (van der Zee et al., 2002). From a content-related point of view, it is possible to explain this delayed predictive effect of individual Conscientiousness by considering an additional perspective on empathy: The ability to understand and to share the emotions of others may be not only regarded as a relational investment (see Section 6.3.2) or as a question of individual skill (see the results on Openness), but also as a question of effort. In other words, empathy may be dependent on the motivation to learn how to understand and how to share the emotions of a specific person. As previous studies documented that individual Conscientiousness and cognitive effort are substantially interrelated (Yeo & Neal, 2004), and that persons with high levels of Conscientiousness have a higher general motivation to learn than persons with low levels of Conscientiousness (Colquitt, LePine, & Noe, 2000), an association between the trait Conscientiousness and the process of becoming more and more able to understand and to share the emotions of others seems reasonable. The delay of this effect is in accordance with the study by Yeo and Neal (2004), who reported that individuals with high levels of Conscientiousness perform worse than individuals with low levels of Conscientiousness when being confronted with an unknown experimental task, because highly conscientious people tend to work very thoroughly and pay great attention to all details of the task, which rather hinders quick success. However, the authors explicitly state that with more practice, *“highly conscientious individuals might surpass those with low conscientiousness”* (Yeo &

Neal, 2004, p.242). The results of the present study support this thesis, at least with regard to empathy. The elaborate intervention program within the second and the third experimental session (see Section 4.2.2 and Section 4.2.3) addressed both cognitive and emotional empathy and offered the possibility to develop and to practice these abilities. This possibility may have been especially attractive for persons with high levels of individual Conscientiousness and a higher motivation to learn (Colquitt et al., 2000).

The substantial associations between the Big Five personality trait Agreeableness and both self-reported and observable empathy (including the IRI subscale Perspective Taking, the IRI subscale Empathic Concern, the CEEQ-GROUP subscale Perspective Taking, the CEEQ-GROUP subscale Empathic Concern, the MET subscale Emotional Empathy, and the DCS group discussion parameter Empathic Feedback) do not lead to new insights, as they are fully in line with various previous studies on the interrelationship between Agreeableness and empathy (e.g., Barrio et al., 2004; Graziano, Habashi, Sheese, & Tobin, 2007; Kraus, Cote, & Keltner, 2010). As the multifaceted structure of Agreeableness with facets as social trust, altruism, compliance, and tender-mindedness (Costa, McCrae, & Dye, 1991) is very similar to the multifaceted structure of empathy (see Section 2.1.2 and Section 2.1.3), a substantial overlap between both constructs is not very surprising and should not be further evaluated at this point.

The DCS group discussion parameter Dominance was primarily associated with the MET subscale Cognitive Empathy and with the DCS group discussion parameter Empathic Statements both in the first and the fourth experimental session. These findings might seem counterintuitive at first glance, as previous research on the functions of empathy (see Section 2.1.3) was not focused on the possibility that feeling into another person might be associated with dominating this person within social interactions. Further, it is possible to explain the latter association between Dominance and Empathic Statements by common method variance (see Section 6.1.4). However, there are several arguments explaining why an association between Dominance and empathy may actually exist. First, Scholl (2013) demonstrated that the communication parameters Dominance and Affiliation are independent from each other. This means that it is possible to be dominant and friendly, or to be dominant and hostile: Scholl (2013, p.22) used the terms “*promotive control*” and “*restrictive control*” to differentiate between these two possibilities to use power. Thus, as

dominance is not associated with prosocial orientation and behavior, the revealed association between dominance and empathy does not contradict the association between empathy and prosocial behavior (see Section 2.1.3). Second, it seems highly reasonable that understanding the emotional states of another person is associated with the possibility to influence and to control this person, as knowledge is advantage, and advantage is power. This specific power may be used both to help but also to harm the other person, in accordance with Scholl's (2013) differentiation. The third argument addresses the phylogenetic perspective on empathy and leadership: A leader with a high ability to understand and to share the emotions of other persons is able to realize the survival values of empathy, e.g. responses to common threats or the coordination of intra- and intergroup behavior (Plutchik, 1987; see Section 2.3.1), to a greater extent than a leader with a low ability to recognize and to understand the emotions of other persons. Interestingly, the association between Dominance and empathy seems to be limited to naturalistic performance and behavioral measures and is not observable with regard to self-report measures of empathy. This phenomenon points to the possibility that socially dominant persons tend to intuitively use their empathic abilities to influence others without reflecting this process consciously.

In the first experimental session, the DCS group discussion parameter Affiliation was negatively correlated with the MET subscale Emotional Empathy and prosocial orientation within the Public Goods Game, and positively correlated with the MET subscale Cognitive Empathy. In the fourth experimental session, Affiliation was positively associated with self-report measures of cognitive empathy (IRI subscale Fantasy, CEEQ-GROUP subscale Perspective Taking) and emotional empathy (IRI subscale Empathic Concern, CEEQ-GROUP subscale Empathic Concern, CEEQ-GROUP subscale Mirroring), with both MET subscales (Cognitive Empathy and Emotional Empathy), and with the DCS group discussion parameter Empathic Statements. This correlation pattern suggests that persons with high levels of individual emotional empathy tend to be less affiliative towards strangers. However, after a short period of group formation, emotional empathy suddenly becomes a highly important positive predictor for affiliative communication. This longitudinal effect strongly supports the argumentation within Section 6.3.2: If empathy is a far-reaching relational investment, highly empathic persons may be especially inhibited to invest in short-term relationships by being affiliative towards strangers. High empathy in combination with high affiliation towards

strangers cannot be functional, because sharing the emotions of every interaction partner must proximately result in exhaustion and ultimately in reduced fitness. The results on the communication parameter Affiliation suggest that the proposed inhibition is limited to emotional empathy, as the MET subscale Cognitive Empathy was positively associated with Affiliation in the first experimental session. The finding that cognitive empathy is a positive predictor for affiliative communication towards strangers can be in turn interpreted together with the finding that affiliative communication towards strangers is negatively correlated with prosocial orientation within the Public Goods Game: Persons with high levels of cognitive empathy and antisocial intentions may be interested in acting affiliative towards strangers, in order to deceive and to exploit them (Epley, Caruso, & Bazerman, 2006). This possibility was already mentioned in Section 6.3.2 and will be further evaluated in Section 6.5.

The associations between empathy and the three demographic characteristics age, gender, and education must be interpreted with caution, as the student sample of the present investigation is not representative (see Section 4.1.2). Nonetheless, it would be inappropriate to discard them, as they may provide valuable insights.

The positive correlations between age and various empathy measures (e.g., the IRI subscale Perspective Taking, the IRI subscale Empathic Concern, the CEEQ-GROUP subscale Empathic Concern, the CEEQ-GROUP subscale Mirroring, the MET subscale Emotional Empathy, and the DCS group discussion parameter Empathic Statements) suggest that adolescents have lower levels of cognitive and emotional empathy than middle-aged persons (age ranged between 19 and 42; $M=27.43$; $SD=5.35$). This linear effect was found both with regard to self-reported and observable empathy. Even if the sample of the present investigation is not representative and even if this finding only reveals age differences and not actual development, it is highly valuable, as empirical evidence on the development of empathy after adolescence is rare and inconsistent (see Section 2.3.2). With regard to the few empirical studies on the development of empathy across adulthood, the reported linear effect contradicts the finding by Grühn et al. (2008), who reported no changes in self-reported empathy across adulthood, and supports the finding by O'Brien et al. (2013), who reported a quadratic interrelationship between age and empathy: In middle-aged adults, the level of cognitive and emotional empathy was higher than in adolescents and older adults.

On the one hand, it is possible that the reported linear increase in empathy between adolescence and middle adulthood and the correspondent finding by O'Brien et al. (2013) represent cohort effects, as both studies used cross-sectional designs, whereas Grühn et al. (2008) conducted a longitudinal study. On the other hand, there are two arguments for why a quadratic interrelationship between age and empathy may actually exist. First, the increase in empathy between adolescence and middle adulthood is in line with the proposed perspective on empathy as a relational investment (see Section 6.3.2), which seems to be inhibited within short-term relationships and to become increasingly important within long-term relationships. At the age of 20, the character of social and especially romantic relationships is more exploratory than at the age of 30, and at the age of 30, the character of social relationships is more exploratory than at the age of 40 (Penke & Asendorpf, 2008; p.1132), indicating that long-term relationships become increasingly important between adolescence and middle adulthood, most probably because of the increasing importance of reproductive goals. Second, the decline in empathy between middle and late adulthood is in line with the Positivity Effect in late adulthood (see Carstensen & Mikels, 2005, for a review): In accordance with the socioemotional selectivity theory, it is functional for older adults to focus on positive memories and emotions in order to maintain an emotional equilibrium, as older adults are able to pursue only few, emotionally relevant personal goals, in contrast to younger and middle-aged adults. High levels of cognitive and emotional empathy may disturb this equilibrium. However, to confirm this hypothesis it is necessary to conduct further longitudinal studies on the development of empathy across adulthood.

With regard to gender differences in empathy, the analyses of the present study indicate that women have higher levels of self-reported individual empathic concern and personal distress, as well as higher levels of empathic feedback within group discussions. However, there were no significant gender differences with regard to any cognitive empathy measure within the present investigation, and also not with regard to prosocial orientation within the Public Goods Game. The finding that women do not outperform men in recognizing and understanding the emotions of others is not in line with several previous investigations (e.g., Baron-Cohen & Wheelwright, 2004; Derntl et al., 2010; Krach et al., 2009; Toussaint & Webb, 2005) and the phylogenetic framework on gender differences in empathy (see Section 2.3.1). However, it is fully in line with two other remarkable studies: A work by Ickes, Gesn, and Graham (2000b) and an investigation by Klein and Hodges (2001).

Both studies demonstrated that gender differences in empathy measures rather reflect gender-role stereotypes than actual gender differences and rather differential motivation than differential ability. Studies reporting gender differences in empathy measures are usually affected by situational cues suggesting to the participants that differences in empathy or emotionality are being measured. Most probably, such cues activate gender-role stereotypes, because gender differences in empathy disappear if these situational cues are absent (Ickes et al., 2000b; out of this reason, any cue on the topic “*emotions*” or “*empathy*” was eliminated within the present study, see Section 4.1.1). This specific effect is usually referred to as “*stereotype threat*” and has already been documented for other abilities, e.g. mathematical skill (Spencer, Steele, & Quinn, 1999). Therefore, in accordance with gender-role stereotypes, women are probably motivated to score high on empathy measures and men are probably motivated to score low on empathy measures, if they know that empathy is being measured, independently from their actual empathic skills. Klein and Hodges (2001) confirmed this assumption, finding no gender differences in cognitive empathy when participants were being offered money in exchange for empathic accuracy. However, considering all mentioned investigations, including neuroscientific approaches (e.g., Derntl et al., 2010), evidence on gender differences in empathy is inconsistent and demands further investigation.

Years of education were interrelated with several self-report measures of empathy (e.g., the IRI subscale Perspective Taking, the IRI subscale Personal Distress, and the MET subscale Emotional Empathy) and with the DCS group discussion parameter Empathic Feedback. These results are in line with previous studies on the interrelationship between academic achievement and emotional intelligence and can be explained by several factors, e.g. by a reciprocal relationship between empathy and reading: Persons with high levels of cognitive and emotional empathy may have higher abilities to understand and to share the emotions of characters within fictional and historical readings and therefore learn better than persons with low levels of cognitive and emotional empathy, as it is easier to memorize emotionally relevant information than emotionally neutral information (see Feshbach & Feshbach, 2009, for a review). However, to investigate if academic achievement is causally determined by empathy it would be necessary to conduct longitudinal studies. It is also possible that empathy is an outcome and not a determinant of academic achievement, or that both constructs are determined by a hidden third variable.

6.4.2 Systemic States and Traits

The systemic variable group climate was positively correlated with the IRI subscales Perspective Taking and Empathic Concern, and the CEEQ-group subscales Perspective Taking and Mental State Perception. Further, group climate in the first experimental session was negatively correlated with the MET subscale Cognitive Empathy. Thus, group climate seems to be positively associated with self-report measures of both cognitive and emotional empathy and to be negatively associated with naturalistic performance measures of cognitive empathy. A positive association between self-reported group climate and self-reported empathy was already revealed within previous investigations (e.g., Johnson et al., 2005), but it is difficult to explain the negative association between group climate and empathic accuracy within the Multifaceted Empathy Test. However, it is important to consider that this association was only found in the first experimental session, and that the level of general positive affect was lower in this session than in the remaining three experimental sessions (see Section 5.1.2, Table 8), most probably because a social interaction with strangers is associated with higher levels of stress than a social interaction with familiar persons (Martin et al., 2015). This interrelationship indicates the possibility that persons with a high ability to understand the emotions of others recognized and reported the general negative affect within the first session more reliably than persons with a low ability to understand the emotions of others. This effect does not only explain the reported negative correlation between group climate in the first session and cognitive empathy but further signifies that it is important to differentiate between self-report, performance, and behavioral measures of empathy.

The systemic parameter group identification was positively correlated with almost all empathy measures of the present study (the IRI subscales Perspective Taking, Fantasy, and Personal Distress, the CEEQ-GROUP subscales Mental State Perception, Empathic Concern, and Mirroring, the MET subscale Emotional Empathy, prosocial orientation within the Public Goods Game, and the DCS group discussion parameter Empathic Feedback). This finding confirms the results of a previous study by Simon et al. (2000), who reported that empathy-motivated helping is highly dependent on the level of group identification. Considering the main analyses on diversity in Section 5.5 and their interpretation in Section 6.3.1, as well as classical theories on the interrelationship between empathy and self-other similarity (Batson et al., 1981; Cialdini; 1997), the association between empathy and group identification

indicates that a perceived self-other overlap may be a predictor of empathy, if it is operationalized as the level of identification with the other person and not as the level of perceived surface-level diversity. This interpretation modifies classical theories on the interrelationship between empathy and perceived diversity and combines them with the phylogenetic argumentation in Section 6.3.1. However, even if shared goals have been identified as an important predictor of group identification in the present study (see Section 6.3.2), future investigations have to reveal additional predictors of group identification, if it is not dependent on perceived surface-level similarity.

The correlations between the subscale Perceived Diversity and the empathy measures of the present investigation were inconsistent. On the one hand, there were positive associations between Perceived Diversity and the IRI subscale Empathic Concern, the IRI subscale Personal Distress, the CEEQ-GROUP subscale Empathic Concern, and the CEEQ-GROUP subscale Mirroring. On the other hand, Perceived Diversity was associated with decreases in prosocial orientation within the Public Goods Game, decreases in systemic empathic concern, decreases in systemic perspective taking, and decreases in the DCS group discussion parameter Empathic Statements. This pattern suggests that perceived surface-level diversity is a positive predictor for static empathy but a negative predictor for the development of empathy. However and more probably, it reflects the process of discovering deep-level dissimilarities within all experimental groups: At the beginning of the experiment, dissimilarities in deep-level characteristics, e.g. personality traits, were not as evident as at the end of the experiment. These deep-level dissimilarities may have resulted in an increased perceived diversity towards the end of the experiment and subsequently in decreased empathy. This explanation is fully in line with the interpretation of the suppression effect in Section 6.3.1.

The subscale Diversity Beliefs in the first experimental session was only associated with an increase in individual empathic concern between the first and the fourth experimental session. However, as diversity beliefs in the fourth experimental session were not associated with any empathy measure, and because there is no theoretical argument or empirical evidence speaking for a direct interrelationship between diversity beliefs and empathy (only a moderating effect with regard to the interrelationship between diversity and empathy, which was investigated in Section 5.6; van Dick et al., 2008), this single correlation most probably represents a random error and will not be discussed any further.

The cognitive representation of a shared goal was positively associated with various empathy measures (e.g., the IRI subscale Empathic Concern, the CEEQ-GROUP subscale Empathic Concern, the CEEQ-GROUP subscale Mental State Perception, the MET subscales Cognitive Empathy and Emotional Empathy, and prosocial orientation within the Public Goods Game). Considering the results in Section 5.5 and their interpretation in Section 6.3.2, it is highly unlikely that the representation of a shared goal elicits empathy, as the experimental manipulation of such a cognitive representation did neither lead to higher levels nor to an increase of genuine empathy within the shared goal condition. Thus, there are two remaining possible explanations for the revealed association. First, it is possible that highly empathic persons have a higher tendency to perceive and to focus on shared goals within a social group than persons with lower levels of empathy. Such an interrelationship seems reasonable, as high levels of cognitive empathy are associated with high levels of cognitive theory of mind (Walter, 2012; see Section 2.2.1): Persons with high levels of cognitive empathy have better capabilities to understand any mental state of another person, even a mental state being rather independent from emotions, as the cognitive representation of a goal. This argumentation is similar to the explanation of the negative correlation between cognitive empathy and group climate within the first experimental session. Second, it is possible to explain the association between empathy and the representation of a shared goal by a hidden third variable influencing both constructs. Such a hidden third variable may be the attitude towards knowledge sharing. This explanation is indirectly supported by a structural equation model by Chow and Chan (2008), who identified an association between social trust, the attitude towards knowledge sharing, and shared goals within a social group.

6.5 On the Possibility of Antisocial Functions of Cognitive Empathy

Antisocial functions of cognitive empathy were already mentioned in Section 2.1.3. Further, a possible association between cognitive empathy and antisocial exploitation was an important argument in explaining the proposed inhibition of empathy in short-term relationships in Section 6.3.2, and possible antisocial functions of cognitive empathy were also used to explain the positive association between cognitive empathy and affiliation

towards strangers on the one hand, and the negative association between affiliation towards strangers and prosocial orientation on the other hand (see Section 6.4.1).

Both arguments are supported by specific partial correlations between two measures of the present study: The MET subscale Cognitive Empathy and prosocial orientation within the Public Goods Game (see Table A.22 in the Appendix, Section 8.3). In the first experimental session, the ability to recognize and to understand the emotional facial expressions of others was positively correlated with prosocial orientation towards group members ($r=.26$; $p<.05$). In the fourth experimental session, the correlation between both measures was not significant anymore ($r=-.13$; n.s.). Further and most interestingly, an increase in cognitive empathy between the first and the fourth experimental session was significantly associated with a decrease in prosocial orientation within the Public Goods Game between the first and the fourth experimental session ($r=-.45$; $p<.001$). This result indicates that participants who enhanced their abilities to recognize and to read the emotions of others after completing the elaborate intervention program addressing cognitive empathy (see Section 4.2.2) used these enhanced abilities to gain advantage within the Public Goods Game and to exploit other group members, as knowing another person's intentions within this game can be easily used to maximize individual profit (see Section 4.3.3).

This result is in line with a study by Epley et al. (2006) carrying the meaningful title *"When Perspective Taking Increases Taking"*. In a series of experiments, the authors revealed that the individual level of cognitive perspective taking predicts reactive egoism in cooperative contexts. A study by Galinsky et al. (2008) replicated this association between perspective taking and the tendency to maximize individual profit on cost of others, and another study by Maddux et al. (2008) revealed an association between emotional mimicry and antisocial exploitation. As both perspective taking and emotional mimicry are substantially interrelated with empathy (see Section 2.2.4 and Section 2.2.5), the three mentioned studies support the presented interpretation of the negative correlation between a naturalistic measure of cognitive empathy and prosocial orientation within the present study.

Nonetheless, antisocial functions of the ability to understand the emotions of others are still rarely addressed within empirical investigations and theoretical models of empathy,

despite their practical relevance and several theoretical arguments speaking for their existence. For example, it seems reasonable that a torturer works most efficiently, if he or she perfectly understands what actions induce the highest level of pain and fear within the tortured person, without directly sharing this pain and fear. This combination of high cognitive empathy and low emotional empathy was actually revealed in individuals with antisocial personality disorder (Blair, 2005) and is immanent within commonsense perspectives on several famous historical figures (e.g., Niccolò Machiavelli, 1469-1527), prominent personalities (e.g., Bernard Madoff, *1938), or even fictional characters (e.g., Frank Underwood from the television series *House of Cards*; Fincher, 2013).

As a matter of course, the proposed association between cognitive empathy and antisocial behavior does not question the various prosocial functions of both cognitive and emotional empathy (see Section 2.1.3). However, it supports the conceptualization of cognitive empathy as a tool, which can be used to realize both prosocial and antisocial intentions. Further, even if cognitive and emotional empathy are usually associated (see Section 2.1.2), it is possible that both dimensions are completely dissociated in some specific cases (e.g., in individuals with an antisocial personality disorder; Blair, 2005). This dissociation may be an important predictor of an antisocial usage of cognitive empathy.

In sum, the results of the present study, the mentioned empirical investigations (Epley et al., 2006; Galinsky et al., 2008; Maddux et al., 2008), and the presented theoretical considerations call for further empirical studies on antisocial functions of cognitive empathy.

6.6 Practical Implications

The results of the present investigation are primarily relevant for the professional practice of clinical, industrial, and educational psychologists.

Several clinical intervention methods are designed to enhance cognitive empathy (e.g., Beck & Feldman, 1989; Charlop-Christy & Daneshvar, 2003; Constanzo, 1992; Delano, 2007; Elfenbein, 2006; Feldman et al., 1992; Gillis et al., 1995) or emotional empathy (Erera, 1997; Herbek & Yammarino, 1990; Pecukonis, 1997; Sherman, 2008; Wastell et al., 2009).

The therapeutic success of these methods is acceptable if they are focused on elementary functions of cognitive empathy as the recognition of emotional facial expressions (e.g., Elfenbein, 2006), but it is limited if they address emotional empathy (Erera, 1997). Further, there are no elaborate intervention programs addressing higher-order functions of empathy as reasoning on another person's emotional state or empathic concern (see Section 2.1.2), except for unspecific therapeutic guidelines (e.g., Block-Lerner, Adair, Plumb, Rhatigan, & Orsillo, 2007, propose "*mindfulness- and acceptance-based behavioral approaches*" to enhance empathic concern). However, the need for successful clinical intervention methods addressing both basic and higher-order functions of either cognitive or emotional empathy is high, because distinct deficits in these functions of empathy were identified in several clinical populations. For example, schizophrenia patients seem to have deficits in basic functions of cognitive empathy (e.g., the recognition of facial expressions, see the meta-analysis by Mandal, Pandey, & Prasad, 1998), and in higher-order functions of cognitive empathy (e.g., reasoning on emotional mental states; see Langdon, Coltheart, & Ward, 2006; Walter et al., 2011), but no deficits in both basic and higher-order functions of emotional empathy (e.g., emotional contagion and empathic concern; see Montag, Heinz, Kunz, & Gallinat, 2007). Several empirical investigations revealed that a similar combination of impaired cognitive empathy and fully intact emotional empathy is observable in persons with autistic spectrum disorders (e.g., Baron-Cohen et al., 1999; Bird et al., 2010; Smith, 2009). Cognitive empathy seems to be impaired in persons with borderline personality disorder, too (Preissler, Dziobek, Ritter, Heekeren, & Roepke, 2010). However, these persons seem to have even higher levels of emotional empathy than average persons, in contrast to the previously mentioned disorders (Harari, Shamay-Tsoory, Ravid, & Levkovitz, 2010). On the other hand, basic and higher-order functions of emotional empathy seem to be impaired in individuals with antisocial personality disorder while functions of cognitive empathy are intact or even above average (Bird & Viding, 2014; Blair, 2005), which was also documented for persons with narcissistic personality disorder (Ritter et al., 2011).

These findings suggest that it is important to design effective and efficient intervention methods addressing the full functional spectrum of either cognitive or emotional empathy. However, the results of the present study specify the difficulties in designing such methods and offer explanations for the limited success of current intervention programs.

Most importantly, the present study points to the possibility that genuine empathy is inhibited within short-term relationships. Despite the elaborate and intensive intervention program, which combined current intervention methods with the objective to enhance cognitive and emotional empathy (see Section 4.2), there were no significant changes in the level of empathy over a period of three weeks, neither in the total sample nor in the three experimental conditions (see Section 5.1 and Section 5.5). This initial inhibition of empathy is reasonable, as the higher-order processes of genuine empathy probably involve a far-reaching relational investment and high levels of commitment (see Section 6.3.2). In this regard, it seems highly questionable if it is at all possible to induce the fundamental willingness to invest that much into a specific relationship by clinical intervention methods, and it seems even more questionable if it is possible to induce a general willingness to invest that much into unspecific relationships. Further, even if this was possible, it is necessary to reconsider the therapeutic desirability of such an intervention, mainly because of possible antisocial functions of cognitive empathy (see Section 6.5) and possible negative effects of emotional empathy, e.g. emotional instability and distress (see Section 6.4.1; LeBlanc et al., 2012). Several theoretical arguments and empirical findings of the present study support the assumption that the initial inhibition of empathy is functional, e.g. the negative correlation between emotional empathy and affiliation towards strangers (see Section 6.4.1). It is even possible that specific impairments in empathy are similarly functional. For example, a limited ability to recognize and to understand the emotions of others may shield persons with autistic spectrum disorders from the negative effects of emotional empathy, if these persons were more sensitive to these negative effects than average persons. This specific thesis is supported by an empirical investigation by Liss, Saulnier, Fein, and Kinsbourne (2006), who documented a striking pattern of overfocused attention, overreactivity, and exceptional memory in persons with autistic spectrum disorders – a combination of traits implying that these persons are highly vulnerable to the negative outcomes of emotional empathy. In accordance with these arguments, it may be oversimplified if not even irresponsible to think that enhancing either cognitive or emotional empathy is an adequate treatment of impairments in these functions without considering the therapeutic risks of such an intervention.

However, the present study does not only point to the limits and the risks of clinical interventions with the objective to enhance empathy but also to an alternative path realizing

prosocial orientation as the most important and often desirable function of empathy (see Section 2.1.3). First, the level of general positive affect probably serves as a heuristic indicator for how risky emotional investments in short-term relationships are, and it is therefore a significant predictor of prosocial orientation within short-term relationships, similarly as the level of group identification (see Section 6.3.2). Second, both mentioned parameters are in turn predicted by the cognitive representation of a shared goal, cooperative orientation, goal commitment, and intrinsic motivation. It is possible to manipulate these four group parameters by a procedure as it was designed and presented in this study (see Section 4.2, 5.2, and 6.2). Third, it is possible that the perception of surface or deep-level dissimilarities between the group members inhibits prosocial orientation (see Section 6.3.1). However, it is necessary to verify the latter assumption by further longitudinal studies measuring indicators of both surface and deep-level diversity.

In sum, any clinical intervention with the objective to elicit prosocial orientation might be more efficient and effective by focusing these group parameters instead of genuine empathy. Other clinical interventions with the objective to enhance social functionality by inducing cognitive or emotional empathy have to consider the associated therapeutic risks, which were pointed out within the present study.

The results of the present investigation are also important for industrial and educational psychologists, as they point to potentially crucial determinants of group climate and group performance. Previous investigations already documented that the perception of a shared goal is associated with group identification (Wegge & Hasslam, 2005), positive affect (Klein et al., 2009), and group performance (Kerr & Tindale, 2004; Locke & Latham, 2002; O'Leary-Kelly et al., 1994; Widmeyer & Ducharme, 1997). Several psychological interventions with the objective to reduce intergroup aggression, stereotypes (e.g., based on the Intergroup Contact Hypothesis, see Pettigrew & Tropp, 2005), and intragroup aggression (e.g., Jigsaw-Techniques, see Aronson, 1978, and Perkins & Saris, 2001) address this interrelationship. These interventions usually enhance group performance and functionality both in industrial and educational contexts. However, the present investigation sheds light on the mechanism behind these interventions: The cognitive representation of a shared goal elicits prosocial orientation and cooperation, and this effect is mediated by general positive affect and moderated by group identification. Furthermore, this study specifies adequate

and efficient techniques to induce a cognitive representation of a shared goal at the beginning of group formation (see Section 4.2, 5.2, and 6.2) and it exemplifies that cooperative orientation, goal commitment, and intrinsic motivation are important variables that have to be additionally addressed in this context. Similarly as within clinical interventions, it might be very difficult, risky, and not necessary to address the higher-order processes of genuine empathy within industrial and educational intervention programs with the objective to enhance group performance, cooperation, and prosocial orientation.

The suppression effect of perceived diversity on the interrelationship between surface-level diversity and prosocial orientation (see Section 6.3.1) indicates that both industrial and educational psychologists have to consider the level of surface and deep-level dissimilarities within the social groups they are working with. Surface and deep-level dissimilarities are associated with benefits as cognitive heterogeneity but also with risks (van Knippenberg et al., 2004), and a reduced prosocial orientation because of perceived dissimilarities might be one of these risks. However, the interrelationship between perceived diversity and group performance is influenced by several parameters, e.g. the meta-contrast ratio (see Section 6.2), diversity beliefs (van Dick et al., 2008), etc. These parameters must also be considered when working with work teams or students.

6.7 Limitations

As within every empirical investigation, the results and implications of this study are limited. The following section will address the most important methodological and content-related limitations.

As already mentioned within Section 6.3.1 and Section 6.3.2, the investigation period of the present study was probably too short. The low stability of most systemic measures indicates that the participants in the experimental groups did not develop stable interaction patterns, mental representations of the other group members, and interaction scripts. In other words, the measures of the present study reflected systemic states rather than systemic traits (Asendorpf & Wilpers, 1998; see Section 2.4.1 and Section 6.1.3), which is a possible explanation for the reduced construct validity of systemic empathy (see Section 6.1.4) and for finding no effects of diversity and shared goals on systemic empathy (see

Section 6.3). Unfortunately, it remains unclear if stable systemic empathy develops in later stages of group formation or if systemic empathy is generally unstable and dependent on affective states (e.g., group climate). Further, it remains unclear if the postulated inhibition of empathy within short-term relationships actually exists (see Section 6.3.2). To verify this assumption, it would be necessary to conduct further empirical investigations on small groups with a longer investigation period than in the present study. According to standard models of small group development (Gersick, 1988; Tuckman, 1965; Tuckman & Jensen, 1977), an investigation period of six or nine weeks would be associated with more valuable insights than a period of only three weeks. However, it is necessary to consider that six weeks would double and that nine weeks would triple the tremendous costs of the present study: 5670,- Euro have been paid to the test persons for their participation (see Section 4.1.1), and 976,- Euro were spent on the Public Goods Game (see Section 4.3.3). A smaller remuneration is not advisable, as it would inevitably result in higher rates of dropout than within the present study (see Section 4.1.1), which have already been critical.

This argument points to the second important limitation of this investigation: Several recruited participants did not attend the first experimental session and seven participants cancelled the experiment before completing the fourth experimental session (see Section 4.1.1). This resulted in different group sizes and most probably in artificially fluctuating group dynamics within several experimental groups throughout the investigation period. Each group member within an experimental group is part of this system and defines it by specific experiences and behaviors. Thus, if one group member cancels the experiment and leaves the group, the measurement of systemic empathy may refer to a completely different system immediately afterwards. However, dropout is a frequent problem within longitudinal studies and even if it can be handled by statistical methods as in the present investigation (see Section 4.4.1 and Hogan, Roy, & Korkontzelou, 2004, for a review), there are no means to prevent it completely. Nonetheless, with more funds and larger sample sizes than in the present study it would be possible to exclude the experimental groups being affected by dropout from all statistical analyses and to investigate groups with unchanging social constellations only.

The results of the present study may be further limited by the circumstance that the experimenters were not blind to the experimental condition. It was not possible to make the experimenters blind to the three experimental conditions, because the experimental

procedure differed between the control condition and the shared goal condition (see Section 4.2), and because the low diversity condition was easily identifiable by a group's composition (e.g., if there were only female participants within an experimental group). However, this limitation was addressed by standardizing all oral instructions and possible answers on questions by the participants within the present study and by extensively training the experimenters.

Another methodological restriction is the student sample of the present study, which cannot be regarded as representative and unbiased, e.g. with regard to important and empathy-related parameters as education, intelligence, age (see Section 6.4.1), systemic meta-contrast ratio (see Section 6.2), etc. Again, as economic considerations motivated the specific recruitment methods within this investigation (see Section 4.1.2), this problem can be easily handled with more funds. This is also the case with regard to the small sample size, which was already addressed throughout Section 4.4.

Moreover, to compute a full multitrait-multimethod model (Eid, 2000) and the exact common-method variance within individual and systemic empathy measures it would have been necessary to include further measures within the present study, e.g. individual communication parameters, systemic image-based empathy measures, etc. Unfortunately, the development and evaluation of these additional measures was far beyond the scope of this work.

In addition to the mentioned methodological limitations, there are several content-related limitations within the present study.

Several empathy-related parameters were not assessed within this investigation, most importantly cultural norms. For example, Cheon et al. (2011) revealed that the cultural preference for social hierarchy is associated with self-reported empathy and with several neurophysiological correlates of empathy (e.g., an increased activity in the left temporo-parietal junction, see Section 2.3.3). An assessment of cultural norms would have been especially important because the experimental groups within the low diversity condition were ethnically homogeneous in contrast to the other two experimental conditions (see Section 4.1.1). Moreover, it would have been important to measure indicators of deep-level diversity and perceived deep-level diversity (see Section 6.3.1). However, as the present study already provides a vast number of control measures (see Section 4.3), there was no

possibility to assess these additional parameters. Further, it was not possible to extend the scope of this work by addressing as comprehensive topics as the interrelationship between cultural norms and empathy.

6.8 Directions for Future Research

First and most importantly, future empirical investigations have to focus on long-term relationships in order to study the emergence and to extrapolate the specific determinants of systemic empathy, even if this focus would be only realizable within a quasi-experimental design. This is the only possibility to verify the postulates of Section 6.3.2: (a) empathy is inhibited within short-term relationships, (b) general positive affect is a significant predictor of prosocial orientation within short-term relationships, and (c) genuine empathy is a significant predictor of prosocial orientation within long-term relationships. The extension of the investigation period within future randomized experimental designs was already mentioned in the previous section.

Second, it would be highly interesting to study the interrelationship between individual empathy, systemic empathy, and further systemic parameters as group cohesion (e.g., Johnson et al., 2005), group performance in various tasks (e.g., Webber & Donahue, 2001), information processing (e.g., Chow & Chan, 2008), deep-level diversity (e.g., Harrison et al., 1998) and various parameters of inter- and intragroup aggression (e.g., Scheithauer, Hayer, Petermann, & Jugert, 2006).

Third, it might be fruitful to measure empathy-related constructs as mentalizing (see Section 2.2.1) or emotional contagion (see Section 2.2.2) together with systemic empathy in order to further specify the differences between them, and in order to verify certain postulates of the present investigation, e.g. that general positive affect is a predictor of emotional contagion but not of genuine empathy within short-term relationships (see Section 6.3.2).

Fourth, future studies may attempt to integrate individual, dyadic, systemic, and intersystemic perspectives on empathy (see Section 2.4.2). To measure these different facets of empathy, it would be necessary to extend current models of social systems (e.g., the latent group model by Gonzales and Griffin, 2002, or latent space models of social networks, Hoff et al., 2002) by evaluating the development of empathy as an individual, dyadic, inter-

dyadic, triadic, inter-triadic, systemic, and intersystemic trait together within one single matrix. The development of such a complex statistical method requires not only a lot of time and effort but also tremendous statistical expertise and ingenious programming skills. However, it would provide insights on the nature of empathy as no empirical method before.

6.9 Main Conclusions and Final Remark

The results and implications of the present study can be most appropriately described by the term ambiguous. The two hypotheses of the present investigation were neither clearly confirmed nor clearly disproved. However, there are three main and practically relevant conclusions. First, it is possible to manipulate group parameters as perceived diversity, the cognitive representation of a shared goal, goal commitment, cooperative orientation, intrinsic motivation, general positive affect, and the level of group identification within a psychological experiment. Second, the cognitive representation of a shared goal elicits systemic positive affect and higher levels of group identification, which in turn elicit supportive and prosocial behavior within social groups. Third, it is very difficult and maybe not possible to manipulate genuine empathy, including higher-order processes as empathic concern and reasoning on another person's emotional states within a psychological experiment. From a phylogenetic perspective, genuine empathy goes beyond emotional contagion and the basic recognition of facial expressions and is most probably a far-reaching relational investment, which may be inhibited in short-term relationships. This restriction has been rather not considered within previous investigations and classical theories on empathy.

In sum, the present investigation challenges and complicates classical perspectives on the ability to understand and to share the emotions of others, e.g. the dependence of empathy on perceived similarity or the empathy-altruism hypothesis. Thus, it is important to conduct further research on the determinants and on the possibility to induce empathy, especially longitudinal studies on empathy within long-term relationships.

However, the ambiguity of the results and the subsequent complications do not imply that this psychological investigation is less valuable and noteworthy than others. The following personal statement will address and explain this point of view.

In 2011, Simmons, Nelson, and Simonsohn published a remarkable article in *Psychological Science*, entitled *“False Positive Psychology”*. Based on several examples, statistical analyses, and two empirical studies, they demonstrate that false-positive findings are dramatically overrepresented in psychological research, proximately because of flexibility in data collection and analysis, and ultimately because *“it is uncommon for prestigious journals to publish null findings”* (Simmons et al., 2011, p.1) and because scientists yield *“to the pressure to do whatever is justifiable to compile a set of studies that we can publish”* (Simmons et al., 2011, p.7). The conclusion of Simmons et al. (2011) is in line with the subjective experience of the author of the present study: strategic considerations seem to be an integral and maybe the most important component of the scientific community’s mindset.

In fact, it would have been very easy to focus on the significant findings within this study, to adjust the theory section and hypotheses to these effects, and to exclude all not significant results, in order to increase the chances of a prestigious publication. Further, it would have been easy to discuss the limitations of the present study (e.g., with regard to the construct validity, see Section 6.1.4) not as open-mindedly and honestly as it was done here. A not uncommon advice with regard to writing scientific papers is to discuss only minor limitations, which can be easily coped with, and to play them down immediately after mentioning them.

But do these strategic considerations really reflect the true nature of science? Are not significant results and not confirmed hypotheses really of lower value than significant results and confirmed hypotheses?

In the eyes of the author, there is only one possible answer on these questions, as the most important task of a scientist cannot be to write fancy stories or to publish in high-ranking journals. The most important task of a scientist is to honestly and sincerely search for the truth, to the very best of his or her knowledge and conscience. Therefore, published papers must not be the *“currency of science”*, as it is often described (Raff, Johnson, & Walter, 2008). As this point of view is not fully correspondent with reality, it may represent an ideal or even a utopian ideal. Nonetheless, this monograph was written with the purpose to realize this idealistic point of view as far as it is possible, and I am grateful that my outstanding advisors Prof. Scheithauer and Prof. von Scheve made this work possible. It was worth all the effort and time, if truth is a little bit closer now.

7. References

- Aiken, L., Stein, J., & Bentler, P. (1994). Structural equation analyses of clinical subpopulation differences and comparative treatment outcomes: Characterizing the daily lives of drug addicts. *Journal of Consulting and Clinical Psychology, 62*, 488-499.
- Allison, P. (2009). Missing data. In R. Millsap & A. Maydeu-Olivares (Eds.), *The sage handbook of quantitative methods in psychology* (pp. 72-89). London: Sage.
- Amabile, T. (1982). The social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology, 43*, 997-1013.
- Andrade, E. & Ariely, D. (2009). The enduring impact of transient emotions on decision making. *Organizational Behavior and Human Decision Processes, 109*, 1-8.
- Armitage, C. & Conner, M. (2001). Efficacy of the theory of planned behavior: A meta-analytic review. *British Journal of Social Psychology, 40*, 471-499.
- Aronson, E. (1978). *The jigsaw classroom*. Beverly Hills, CA: Sage.
- Asendorpf, J. & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology, 74*, 1531-1544.
- Asendorpf, J. & Banse, R. (2000). *Psychologie der Beziehung. [Relationship psychology]*. Bern: Huber.
- Asendorpf, J. & Neyer, F. (2013). *Psychologie der Persönlichkeit. [Personality psychology]*. Berlin: Springer.
- Astington, J., Harris, P., & Olson, D. (1988). *Developing theories of mind*. New York: Cambridge University Press.
- Aubé, C. & Rousseau, V. (2005). Team goal commitment and team effectiveness: The role of task interdependence and supportive behaviors. *Group Dynamics: Theory, Research, and Practice, 9*, 189-204.
- Austin, J. & Vancouver, J. (1996). Goals constructs in psychology: Structure, process, and content. *Psychological Bulletin, 120*, 338-375.
- Baaren, R. v., Holland, R., Kawakami, K., & Knippenberg, A. v. (2004). Mimicry and prosocial behavior. *Psychological Science, 15*, 71-74.
- Baas, M., De Dreu, C., & Nijstad, B. (2008). A meta-analysis of 25 years of mood-creativity research: Hedonic tone, activation, or regulatory focus? *Psychological Bulletin, 134*, 779-806.
- Baer, J. (1996). Does artistic creativity decline during elementary school? *Psychological Reports, 78*, 927-930.
- Bailey, P., Henry, J., & von Hippel, W. (2008). Empathy and social functioning in late adulthood. *Aging and Mental Health, 12*, 499-503.

- Baldwin, M. (1992). Relational schemas and the processing of social information. *Psychological Bulletin*, 112, 461-484.
- Barnett, M. (1987). Empathy and related responses in children. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development*. New York: Cambridge University Press.
- Barnett, V. & Lewis, T. (1994). *Outliers in statistical data*. New York: Wiley.
- Baron-Cohen, S., Ring, H., Wheelwright, S., Bullmore, E., Brammer, M., Simmons, A., & Williams, S. (1999). Social intelligence in the normal and autistic brain: An fMRI study. *European Journal of Neuroscience*, 11, 1891-1898.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test Revised Versions: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychiatry and Psychiatry*, 42, 241-252.
- Baron-Cohen, S. & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34, 163-175.
- Barrio, V., Aluja, A., & Garcia, L. (2004). Relationship between empathy and the Big Five personality traits in a sample of Spanish adolescents. *Social Behavior and Personality*, 32, 677-682.
- Barsade, S. & Gibson, D. (1998). Group emotion: A view from top and bottom. *Research on Managing Groups and Teams*, 1, 81-102.
- Barsade, S. (2002). The Ripple Effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly*, 47, 644-675.
- Bastiaansen, J., Thioux, M., & Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical Transactions of the Royal Society B*, 364, 2391-2404.
- Batson, C., Duncan, B., Ackerman, P., Buckley, T., & Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology*, 40, 290-302.
- Batson, C., O'Quin, K., Fultz, J., Vanderplas, M., & Isen, A. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of Personality and Social Psychology*, 45, 706-718.
- Batson, C., Dyck, J., Brandt, J., Batson, J., Powell, A., McMaster, M., & Griffitt, C. (1988). Five studies testing two new egoistic alternatives to the Empathy-Altruism Hypothesis. *Journal of Personality and Social Psychology*, 55, 52-77.
- Batson, C., Batson, J., Slingsby, J., Harrell, K., Peekna, H., & Todd, R. (1991). Empathic joy and the empathy-altruism hypothesis. *Journal of Personality and Social Psychology*, 61, 413-426.
- Batson, C., Early, S., & Salvarani, G. (1997). Perspective taking: Imagining how another feels versus imagining how you would feel. *Personality and Social Psychology Bulletin*, 23, 751-758.

- Batson, C. (2009). These things called empathy: Eight related but distinct phenomena. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 3-15). Cambridge: MIT Press.
- Batson, C. (2011). *Altruism in humans*. New York: Oxford University Press.
- Bavelas, J., Black, A., Lemery, C., & Mullett, J. (1986). I show you how you feel: Motor mimicry as a communicative act. *Journal of Personality and Social Psychology*, 50, 322-329.
- Beck, L. & Feldman, R. (1989). Enhancing children's decoding of facial expression. *Journal of Nonverbal Behavior*, 13, 269-278.
- Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain*, 133, 1515-1535.
- Bird, G. & Viding, E. (2014). The self to other model of empathy: Providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience & Biobehavioral Reviews*, 47, 520-532.
- Bischof-Köhler, D. (1991). The development of empathy in infants. In M. Lamb & H. Keller (Eds.), *Infant development: Perspectives from German-speaking countries*. New Jersey: Erlbaum.
- Blair, R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness & Cognition*, 14, 698-718.
- Blanz, M. (1999). Accessibility and fit as determinants of the salience of social categorizations. *European Journal of Social Psychology*, 29, 43-74.
- Block-Lerner, J., Adair, C., Plumb, J., Rhatigan, D., & Orsillo, S. (2007). The case of mindfulness-based approaches in the cultivation of empathy: Does nonjudgmental, present-moment awareness increase capacity for perspective-taking and empathic concern? *Journal of Marital and Family Therapy*, 33, 501-516.
- Borkenau, P. & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI)*. [NEO-Five-Factor Inventory]. Göttingen: Hogrefe.
- Box, G. & Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, 26, 211-252.
- Brackett, M., Rivers, S., Shiffman, S., Lerner, N., & Salovey, P. (2006). Relating emotional abilities to social functioning: A comparison of self-report and performance measures of emotional intelligence. *Journal of Personality and Social Psychology*, 91, 780-795.
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 32, 513-531.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Bronfenbrenner, U. (1986). The ecology of the family as a context for human development: Research perspectives. *Developmental Psychology*, 22, 723-742.

- Bronfenbrenner, U. & Morris, P. (1998). The ecology of developmental processes. In W. Damon & R. Lerner (Eds.), *Handbook of child psychology, volume 1: Theoretical models of human development* (pp. 993-1028). New York: John Wiley & Sons.
- Brown, M. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. Bollen & J. Long (Eds.), *Testing structural equation models* (pp. 445-455). Newbury Park: Sage.
- Brown, T. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford.
- Campbell, D. & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Capra, C. (2004). Mood-driven behavior in strategic interactions. *The American Economic Review*, 94, 367-372.
- Carmines, E. & Stimson, J. (1980). The two faces of issue voting. *The American Political Science Review*, 74, 78-91.
- Carr, J., Schmidt, A., Ford, K., & DeShon, R. (2003). Climate perceptions matter: A meta-analytic path analysis relating molar climate, cognitive and affective states, and individual level work outcomes. *Journal of Applied Psychology*, 88, 605-619.
- Carstensen, L. & Mikels, J. (2005). At the intersection of emotion and cognition: Aging and the Positivity Effect. *Current Directions in Psychological Science*, 14, 117-121.
- Charlop-Christy, M. & Daneshvar, S. (2003). Using video modeling to teach perspective taking to children. *Journal of Positive Behavior Interventions*, 5, 12-21.
- Cheon, B., Im, D., Harada, T., Kim, J., Mathur, V., Scimeca, J., Parrish, T., et al. (2011). Cultural influences on neural basis of intergroup empathy. *NeuroImage*, 57, 642-650.
- Choi, J., Price, R., & Vinokur, A. (2003). Self-efficacy changes in groups: Effects of diversity, leadership, and group climate. *Journal of Organizational Behavior*, 24, 1-16.
- Chow, W. & Chan, L. (2008). Social network, social trust and shared goals in organizational knowledge sharing. *Information & Management*, 45, 458-465.
- Cialdini, R., Brown, S., Lewis, B., Luce, C., & Neuberg, S. (1997). Reinterpreting the empathy-altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology*, 73, 481-494.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum.
- Cohen, D. & Strayer, J. (1996). Empathy in conduct-disordered and comparison youth. *Developmental Psychology*, 32, 988-998.
- Colquitt, J., LePine, J., & Noe, R. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology*, 85, 678-707.
- Constanzo, M. (1992). Training students to decode verbal and nonverbal cues: Effects on confidence and performance. *Journal of Nonverbal Behavior*, 23, 91-100.

- Cook, W. & Kenny, D. (2005). The Actor-Partner Interdependence Model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development, 29*, 101-109.
- Cooper, R. & Jayatilaka, B. (2006). Group creativity: The effects of extrinsic, intrinsic, and obligation motivations. *Creativity Research Journal, 18*, 153-172.
- Corey, D., Dunlap, W., & Burke, M. (1998). Averaging correlations: Expected values and bias in combined Pearson *r*s and Fisher's *z* transformations. *The Journal of General Psychology, 125*, 245-261.
- Corey, G. (2012). *Theory and practice of group counseling*. Belmont, CA: Brooks.
- Costa, P., McCrae, R., & Dye, D. (1991). Facet scales for Agreeableness and Conscientiousness: A revision of the Neo Personality Inventory. *Personality and Individual Differences, 9*, 887-898.
- Costa, P. & McCrae, R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Odessa: Psychological Assessment Resources.
- Cronbach, L. (1951). Coefficient Alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.
- Crowder, M. & Hand, D. (1990). *Analysis of repeated measures*. London: Chapman & Hall.
- Curran, S., Andrykowski, M., & Studts, J. (1995). Short form of the Profile of Mood States (POMS-SF): Psychometric information. *Psychological Assessment, 7*, 80-83.
- Davis, M. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85.
- Davis, M. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113-126.
- Davis, M. & Franzoi, S. (1991). Stability and change in adolescent self-consciousness and empathy. *Journal of Research in Personality, 25*, 70-87.
- Davis, M. (1994). *Empathy: A social psychological approach*. Boulder, CO: Westview Press.
- de Vignemont, F. & Singer, T. (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences, 10*, 435-441.
- de Waal, F. (2004). On the possibility of animal empathy. In A. Manstead, N. Frijda, & A. Fischer (Eds.), *Feelings and emotions* (pp. 381-398). New York: Cambridge University Press.
- de Waal, F. (2008). Putting the altruism back into the altruism: The evolution of empathy. *Annual Review of Psychology, 59*, 279-300.
- Decety, J. & Sommerville, J. (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences, 7*, 527-533.
- Decety, J. & Jackson, P. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*, 71-100.

- Decety, J. & Hodges, S. (2006). The social neuroscience of empathy. In P.A.M. van Lange (Ed.), *Bridging social psychology: Benefits of transdisciplinary approaches*. Mahwah, NJ: Erlbaum.
- Decety, J. & Jackson, P. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science*, 15, 54-58.
- Delano, M. (2007). Video modeling interventions for individuals with autism. *Remedial and Special Education*, 28, 33-42.
- Derntl, B., Finkelmeyer, A., Eickhoff, S., Kellermann, T., Falkenberg, D., Schneider, F., & Habel, U. (2010). Multidimensional assessment of empathic abilities: Neural correlates and gender differences. *Psychoneuroendocrinology*, 35, 67-82.
- Dinstein, I., Gardner, J., Jazayeri, M., & Heeger, D. (2008). Executed and observed movements have different distributed representations in human aIPS. *Journal of Neuroscience*, 28, 11231-11239.
- Doherty, W., Orimoto, L., Singelis, T., Hatfield, E., & Hebb, J. (1995). Emotional contagion: Gender and occupational differences. *Psychology of Women Quarterly*, 19, 355-371.
- Dunbar, R. (1998). The social brain hypothesis. *Evolutionary Anthropology*, 6, 178-190.
- Dunn, D. (2011). Situations matter: Teaching the Lewinian link between social psychology and rehabilitation psychology. *History of Psychology*, 14, 405-411.
- Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., Kessler, J., et al. (2006). Introducing MASC: A movie for the assessment of social cognition. *Journal of Autism and Developmental Disorders*, 36, 623-636.
- Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H., Wolf, O., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders*, 38, 464-473.
- Dziobek, I. (2012). Towards a more ecologically valid assessment of empathy. *Emotion Review*, 4, 18-19.
- Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika*, 65, 241-261.
- Eisenberg, N. & Strayer, J. (1987). *Empathy and its development*. New York: Cambridge University Press.
- Eisenberg, N., Fabes, R., Miller, P., Fultz, J., Shell, R., Mathy, R., & Reno, R. (1989). Relation of sympathy and personal distress to prosocial behavior: A multimethod study. *Journal of Personality and Social Psychology*, 57, 55-66.
- Eisenberg, N. & Fabes, R. (1990). Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, 14, 131-149.

- Eisenberg, N. (2000). Empathy and sympathy. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of emotion* (pp. 677-691). New York: Guilford.
- Eisenberg, N. (2002). Distinctions among various modes of empathy-related reactions: A matter of importance in humans. *Behavioral and Brain Sciences*, 25, 33-34.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169-200.
- Elfenbein, H. (2006). Learning in emotion judgments: Training and the cross-cultural understanding of facial expressions. *Journal of Nonverbal Behavior*, 30, 21-36.
- Epley, N., Caruso, E., & Bazerman, M. (2006). When perspective taking increases taking: Reactive egoism in social interaction. *Journal of Personality and Social Psychology*, 91, 872-889.
- Erera, P. (1997). Empathy training for helping professionals: Model and evaluation. *Journal of Social Work Education*, 33, 245-260.
- Federal Ministry for Family, Senior Citizens, Women, and Youth (2005). Familienfreundlichkeit im Betrieb: Ergebnisse einer repräsentativen Umfrage. [Family-friendliness at the workplace: Results of a representative survey.] Retrieved from: http://www.ifd-allensbach.de/uploads/tx_studies/7036_Familienfreundlichkeit.pdf on 01/03/2013.
- Fehr, E. & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90, 980-994.
- Fehr, E. & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415, 137-140.
- Feldman, R., Philippot, P., & Custrini, R. (1991). Social competence and nonverbal behavior. In R. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 329-350). New York: Cambridge University Press.
- Feshbach, N. & Feshbach, S. (2009). Empathy and education. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 85-97). Cambridge: MIT Press.
- Fincher, D. (2013). *House of cards* [Television series]. Los Gatos, CA: Netflix.
- Fisher, R. (1915). Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population. *Biometrika*, 10, 507-521.
- Fisher, R. (1935). The logic of inductive inference. *Journal of the Royal Statistical Society*, 98, 39-54.
- Fiske, S. & Russell, A. (2010). Cognitive processes. In J. Dovidio, M. Hewstone, P. Glick, & V. Esses (Eds.), *The sage handbook of prejudice, stereotyping and discrimination* (pp. 115-130). New York: Sage.
- Forgas, J. (1999). Feeling and speaking: Mood effects on verbal communication strategies. *Personality and Social Psychology Bulletin*, 25, 850-863.
- Galinsky, A., Maddux, W., Gilin, D., & White, J. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective-taking and empathy in negotiations. *Psychological Science*, 19, 378-384.

- Gallup, G. & Platek, S. (2002). Cognitive empathy presupposes self-awareness: Evidence from phylogeny, ontogeny, neuropsychology, and mental illness. *Behavioral and Brain Sciences*, 25, 36-37.
- Gersick, C. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal*, 31, 9-41.
- Gianino, A. & Tronick, E. (1988). The mutual regulation model: The infant's self and interactive regulation and coping and defensive capacities. In T. Field, P. McCabe, & N. Schneiderman (Eds.), *Stress and coping across development* (pp. 47-68). New Jersey: Erlbaum.
- Gieser, T. (2008). Embodiment, emotion and empathy: A phenomenological approach to apprenticeship learning. *Anthropological Theory*, 8, 299-318.
- Gilbert, G. (1961). *Nuremberg diary*. New York: Signet Book.
- Gilbert, P. (2005). *Compassion: Conceptualizations, research and use in psychotherapy*. New York: Routledge.
- Gillis, J., Bernieri, F., & Wooten, E. (1995). The effects of stimulus medium and feedback on the judgment of rapport. *Organizational Behavior and Human Decision Processes*, 63, 33-45.
- Gladstein, G. (1983). Understanding empathy: Integrating counseling, developmental, and social psychology perspectives. *Journal of Counseling Psychology*, 30, 467-482.
- Gonzales, R. & Griffin, D. (1996). Dyadic data analysis: New statistical tools for analyzing relationship data. *International Society for the Study of Personal Relationships Bulletin*, 12, 7-9.
- Gonzales, R. & Griffin, D. (1997). The statistics of interdependence: Treating dyadic data with respect. In S.W. Duck (Ed.), *Handbook of personal relationships: Theory, research, and interventions* (pp. 271-302). New York: Wiley.
- Gonzales, R. & Griffin, D. (2002). Modeling the personality of dyads and groups. *Journal of Personality*, 70, 901-924.
- Graziano, W., Habashi, M., Sheese, B., & Tobin, R. (2007). Agreeableness, empathy, and helping: A person x situation perspective. *Journal of Personality and Social Psychology*, 93, 583-599.
- Greenhouse, S. & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95-112.
- Greimel, E., Schulte-Rüther, M., Fink, G., Piefke, M., Herpertz-Dahlmann, B., & Konrad, K. (2010). Development of neural correlates of empathy from childhood to early adulthood: An fMRI study in boys and adult men. *Journal of Neural Transmission*, 117, 781-791.
- Gross, J. & Levenson, R. (1995). Emotion elicitation using films. *Cognition and Emotion*, 9, 87-108.
- Grühn, D., Rebucal, K., Diehl, M., Lumley, M., & Labouvie-Vief, G. (2008). Empathy across the adult lifespan: Longitudinal and experience-sampling findings. *Emotion*, 8, 753-765.

- Gutsell, J. & Inzlicht, M. (2010). Empathy constrained: Prejudice predicts mental simulation of actions during observation of outgroups. *Journal of Experimental Social Psychology, 46*, 841-845.
- Haas, J., Sypher, B., & Sypher, H. (1992). Do shared goals really make a difference? *Management Communication Quarterly, 6*, 166-179.
- Han, B. (2008). *A typology of empathic communication: Reviewing empathy as a communicative act*. Paper presented at the annual conference of the International Communication Association, Montréal, Canada. Retrieved from: http://citation.allacademic.com/meta/p_mla_apa_research_citation/2/3/4/6/9/pages234694/p234694-1.php on 01/03/2013.
- Harari, H., Shamay-Tsoory, S., Ravid, M., & Levkovitz, Y. (2010). Double dissociation between cognitive and affective empathy in borderline personality disorder. *Psychiatry Research, 175*, 277-279.
- Harrison, D., Price, K., & Bell, M. (1998). Beyond relational demography: Time and the effects of surface-and deep-level diversity on work group cohesion. *The Academy of Management Journal, 41*, 96-107.
- Hatfield, E., Cacioppo, J., & Rapson, R. (1993). Emotional contagion. *Current Directions in Psychological Science, 2*, 96-99.
- Hatfield, E., Rapson, R., & Le, Y. (2009). Emotional contagion and empathy. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 19-30). Cambridge: MIT Press.
- Herbek, T. & Yammarino, F. (1990). Empathy training for hospital staff nurses. *Group & Organization Management, 15*, 279-295.
- Hersch, B. (1990). *TABU*. Pawtucket: Hasbro.
- Heywood, H. (1931). On finite sequences of real numbers. *Proceedings of the Royal Society of London, 134*, 486-501.
- Hirata, S. (2009). Chimpanzee social intelligence: Selfishness, altruism, and the mother-infant bond. *Primates, 50*, 3-11.
- Hoff, P., Raftery, A., & Handcock, M. (2002). Latent space approaches to social network analyses. *Journal of the American Statistical Association, 97*, 1090-1098.
- Hoffmann, M. (2000). *Empathy and moral development: Implications for caring and justice*. New York: Cambridge University Press.
- Hoffmann, M. (2002). How automatic and representational is empathy, and why. *Behavioral and Brain Sciences, 25*, 38-39.
- Hogan, J., Roy, J., & Korkontzelou, C. (2004). Handling drop-out in longitudinal studies. *Statistics in Medicine, 23*, 1455-1497.
- Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology, 33*, 307-316.

- Holt, C. & Laury, S. (1997). Classroom games: Voluntary provision of a public good. *The Journal of Economic Perspectives*, 11, 209-215.
- Howard, G. & Dailey, P. (1979). Response-shift bias: A source of contamination of self-report measures. *Journal of Applied Psychology*, 64, 144-150.
- Hu, L. & Bentler, P. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3, 424-453.
- Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, 59, 730-742.
- Ickes, W., Buysse, A., Pham, H., Rivers, K., Erickson, J., Hancock, M., Kelleher, J., & Gesn, P. (2000a). On the difficulty of distinguishing "good" and "poor" perceivers: A social relations analysis of empathic accuracy data. *Personal Relationships*, 7, 219-234.
- Ickes, W., Gesn, P., & Graham, T. (2000b). Gender differences in empathic accuracy: Differential ability or differential motivation? *Personal Relationships*, 7, 95-109.
- Institute for Public Opinion Research Allensbach (2008). Rauchverbote und Raucher. [Smoking bans and smokers]. Retrieved from: http://www.ifd-allensbach.de/uploads/tx_reportsndocs/prd_0801_01.pdf on 01/03/2013.
- Isen, A. & Lewin, P. (1972). Effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology*, 21, 384-388.
- Jacobs, A. (2012). Comment on Walter's "Social cognitive neuroscience of empathy: Concepts, circuits, and genes." *Emotion Review*, 4, 20-21.
- Jelen, T. & Wilcox, C. (2003). Causes and consequences of public attitudes towards abortion: A review and research agenda. *Political Research Quarterly*, 56, 489-500.
- Johnson, J., Burlingame, G., Olsen, J., Davies, R., & Gleave, R. (2005). Group climate, cohesion, alliance, and empathy in group psychotherapy: Multilevel structural equation models. *Journal of Counseling Psychology*, 52, 310-321.
- Jurasova, K. & Spajdel, M. (2013). Development and assessment of film excerpts used for emotion elicitation. *Activitas Nervosa Superior Rediviva*, 55, 135-140.
- Kalbe, E., Schlegel, M., Sack, A., Nowak, D., Dafotakis, M., Bangard, C., Brand, M., et al. (2010). Dissociating cognitive from affective theory of mind: A TMS study. *Cortex*, 46, 769-780.
- Kelly, J. & Barsade, S. (2001). Mood and emotions in small groups and work teams. *Organizational Behavior and Human Decision Processes*, 86, 99-130.
- Kelly, J. & Spoor, J. (2006). Affective influence in groups. In J. Forgas (Ed.), *Affect in social thinking and behavior* (pp. 311-325). New York: Taylor & Francis.
- Kerr, N. & Tindale, R. (2004). Group performance and decision making. *Annual Review of Psychology*, 55, 623-655.

- Keysers, C. & Perrett, D. (2004). Demystifying social cognition: A Hebbian perspective. *Trends in Cognitive Science*, 8, 501-507.
- Keyton, J. & Beck, S. (2009). The influential role of relational messages in group interaction. *Group Dynamics*, 13, 14-30.
- Kirschner, S. & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, 102, 299-314.
- Klein, C., DiazGranados, D., Salas, E., Le, H., Burke, C., Lyons, R., & Goodwin, G. (2009). Does team building work? *Small Group Research*, 40, 181-222.
- Klein, H., Wesson, M., Hollenbeck, J., Wright, P., & DeShon, R. (2001). The assessment of goal commitment: A measurement model meta-analysis. *Organizational Behavior and Human Decision Processes*, 85, 32-55.
- Klein, K. & Hodges, S. (2001). Gender differences, motivation, and empathic accuracy: When it pays to understand. *Personality and Social Psychology Bulletin*, 27, 720-730.
- Kolenikov, S. & Bollen, K. (2012). Testing negative error variances: Is a Heywood Case a symptom of misspecification? *Sociological Methods & Research*, 41, 124-167.
- Konstantopoulos, S. (2009). Using power tables to compute statistical power in multilevel experimental designs. *Practical Assessment, Research & Evaluation*, 14. Retrieved from: <http://pareonline.net/pdf/v14n10.pdf> on 01/03/2013.
- Krach, S., Blümel, I., Marjoram, D., Lataster, T., Krabbendam, L., Weber, J., van Os, J., et al. (2009). Are women better mindreaders? Sex differences in neural correlates of mentalizing detected with functional MRI. *BMC Neuroscience*, 10, doi:10.1186/1471-2202-10-9
- Kraus, M., Cote, S., & Keltner, D. (2010). Social class, contextualism, and empathic accuracy. *Psychological Science*, 20, 1-8.
- Krebs, D. (1975). Empathy and altruism. *Journal of Personality and Social Psychology*, 32, 1134-1146.
- Krosnick, J., Narayan, S., & Smith, W. (1996). Satisficing in surveys: Initial evidence. *New Directions for Evaluation*, 70, 29-44.
- Krosnick, J. (1999). Survey research. *Annual Review of Psychology*, 50, 537-567.
- Koski, S. & Sterck, E. (2010). Empathic chimpanzees: A proposal of the levels of emotional and cognitive processing in chimpanzee empathy. *European Journal of Developmental Psychology*, 7, 38-66.
- Lamm, C., Batson, C., & Decety, J. (2007). The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience*, 19, 42-58.
- Lang, P., Bradley, M., & Cuthbert, B. (2008). *International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-8. University of Florida, Gainesville.

- Langdon, R., Coltheart, M., & Ward, P. (2006). Empathetic perspective-taking is impaired in schizophrenia: Evidence from a study of emotion attribution and theory of mind. *Cognitive Neuropsychiatry*, 11, 133-155.
- LaPiere, R. (1934). Attitudes versus action. *Social Forces*, 13, 230-237.
- Lawler, E., Thye, S., & Yoon, J. (2000). Emotion and group cohesion in productive exchange. *American Journal of Sociology*, 106, 616-657.
- LeBlanc, D., Gilin, D., Calnan, K., & Solarz, B. (2012). *Perspective taking, empathy, and relational conflict at work: An investigation among participants in a workplace conflict resolution program*. Paper presented at the 25th Annual International Association of Conflict Management Conference, Spier, South Africa, July 12-14. Retrieved from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2086578 on 30/08/2012.
- Lennon, R. & Eisenberg, N. (1987). Gender and age differences in empathy and sympathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 195-217). New York: Cambridge University Press.
- Leslie, K., Johnson-Frey, S., & Grafton, S. (2004). Functional imaging of face and hand imitation: Towards a motor theory of empathy. *NeuroImage*, 24, 601-607.
- Leudar, I., Costall, A., & Francis, D. (2004). Theory of mind: A critical assessment. *Theory & Psychology*, 14, 571-578.
- Levene, H. (1960). Robust tests for equality of variances. In I. Olkin, S. Ghurye, W. Hoeffding, W. Madow, & H. Mann (Eds.), *Contributions to probability and statistics* (pp. 278-292). Stanford: Stanford University Press.
- Levine, J. & Moreland, R. (1990). Progress in small group research. *Annual Review of Psychology*, 41, 585-634.
- Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw Hill.
- Lewin, K., Lipitt, R., & White, R. (1939). Patterns of aggressive behavior in experimentally created "social climates". *The Journal of Social Psychology*, 10, 269-299.
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49, 764-766.
- Liebal, K., Vaish, A., Haun, D., & Tomasello, M. (2014). Does sympathy motivate prosocial behavior in great apes? *Plos One*, 9, e84299.
- Lingnau, A., Gesierich, B., & Caramazza, A. (2009). Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Sciences*, 106, 9925-9930.

- Lippitt, R. (1939). Field theory and experiment in social psychology: Autocratic and democratic group atmospheres. *American Journal of Sociology*, 45, 26-49.
- Liss, M., Saulnier, C., Fein, D., & Kinsbourne, M. (2006). Sensory and attention abnormalities in autistic spectrum disorders. *Autism*, 10, 155-172.
- Little, R. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83, 1998-1202.
- Little, R. & Rubin, D. (2002). *Statistical analysis with missing data*. New York: Wiley.
- Locke, E. & Latham, G. (2002). Building a practically useful theory of goal setting and task motivation. *American Psychologist*, 57, 705-717.
- Loersch, C., Aarts, H., Payne, K., & Jefferis, V. (2008). The influence of social groups on goal contagion. *Journal of Experimental Social Psychology*, 44, 1555-1558.
- Lovett, B. & Sheffield, R. (2007). Affective empathy deficits in aggressive children and adolescents: A critical review. *Clinical Psychology Review*, 27, 1-13.
- Luxen, M. (2005). Gender differences in dominance and affiliation during a demanding interaction. *The Journal of Psychology*, 139, 331-347.
- Maas, C. & Hox, J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, 1, 86-92.
- Maasen, G. & Bakker, A. (2001). Suppressor variables in path models. *Sociological Methods & Research*, 30, 241-270.
- Mackie, D., Devos, T., & Smith, E. (2000). Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology*, 79, 602-616.
- Maddux, W., Mullen, E., & Galinsky, A. (2008). Chameleons bake bigger pies and take bigger pieces: Strategic behavioural mimicry facilitates negotiation outcomes. *Journal of Experimental Social Psychology*, 44, 461-468.
- Maitner, A., Mackie, D., & Smith, E. (2006). Evidence for the regulatory function of intergroup emotion: Consequences of implemented or impeded intergroup action tendencies. *Journal of Experimental Social Psychology*, 42, 720-728.
- Mandal, M., Pandey, R., & Prasad, A. (1998). Facial expressions of emotions and schizophrenia: A review. *Schizophrenia Bulletin*, 24, 399-412.
- Martin, L., Hathaway, G., Isbester, K., Mirali, S., Acland, E., Niederstrasser, N., Slepian, P., et al. (2015). Reducing social stress elicits emotional contagion of pain in mouse and human strangers. *Current Biology*, 25, 1-7.
- Martinovski, B., Traum, D., & Marsella, S. (2007). Rejection of empathy in negotiation. *Group Decision and Negotiation*, 16, 61-76.
- Mason, P. & Barta, I. (2010). How the social brain experiences empathy: Summary of a gathering. *Social Neuroscience*, 5, 252-256.

- Massey, F. (1951). The Kolmogorov-Smirnov Test for goodness of fit. *Journal of the American Statistical Association*, 46, 68-78.
- Mathieu, J., Maynard, M., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of Management*, 34, 410-476.
- Matsui, T., Kakuyama, T., & Onglatco, M. (1987). Effects of goals and feedback on performance in groups. *Journal of Applied Psychology*, 72, 407-415.
- Matthews, G., Campbell, S., Falconer, S., Joyner, L., Huggins, J., Gilliland, K., Grier, R., & Warm, J. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion*, 2, 315-340.
- Mauchly, J. (1940). Significance test for sphericity of a normal n-variate distribution. *The Annals of Mathematical Statistics*, 11, 204-209.
- McCrae, R. & Costa, P. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90.
- McCrae, R. (1994). Openness to experience: Expanding the boundaries of Factor V. *European Journal of Personality*, 8, 251-272.
- McGrath, J., Arrow, H., & Berdahl, J. (2000). The study of groups: Past, present, and future. *Personality and Social Psychology Review*, 4, 95-105.
- Mella, N., Studer, J., Gilet, A., & Labouvie-Vief, G. (2012). Empathy for pain from adolescence through adulthood: An event-related brain potential study. *Frontiers in Psychology*, 3, Article 501.
- Mencl, J. & May, D. (2009). The effects of proximity and empathy on ethical decision-making: An exploratory investigation. *Journal of Business Ethics*, 85, 201-226.
- Meyer, B., Shemla, M., & Schermuly, C. (2011). Social category salience moderates the effect of diversity faultlines on information elaboration. *Small Group Research*, 42, 257-282.
- Mitchell, T. & Silver, W. (1990). Individual and group goals when workers are independent: Effects on task strategies and performance. *Journal of Applied Psychology*, 75, 185-193.
- Montag, C., Heinz, A., Kunz, D., & Gallinat, J. (2007). Self-reported empathic abilities in schizophrenia. *Schizophrenia Research*, 92, 85-89.
- Moody, J. & White, D. (2003). Structural cohesion and embeddedness: A hierarchical concept of social groups. *American Sociological Review*, 68, 103-127.
- Moutafi, J., Furnham, A., & Crump, J. (2006). What facets of openness and conscientiousness predict fluid intelligence score? *Learning and Individual Differences*, 16, 31-42.
- Muccino, G. (2006). *The pursuit of happiness* [Motion picture]. Los Angeles: Sony Pictures.
- Mulvey, P. & Ribbens, B. (1999). The effects of intergroup competition and assigned group goals on group efficacy and group effectiveness. *Small Group Research*, 30, 651-677.
- Myers, D. & Diener, E. (1995). Who is happy? *Psychological Science*, 6, 10-19.

- Nadler, A. & Liviatan, I. (2006). Intergroup reconciliation: Effects of adversary's expressions of empathy, responsibility, and recipients' trust. *Personality and Social Psychology Bulletin*, 32, 459-470.
- Neyer, F. & Lang, F. (2003). Blood is thicker than water: Kinship orientation across adulthood. *Journal of Personality and Social Psychology*, 84, 310-321.
- Nowak, A. (2011). *Introducing a pedagogy of empathic action as informed by social entrepreneurs*. (Doctoral Dissertation, McGill University, Canada). Retrieved from: http://digitool.library.mcgill.ca/R/?func=dbin-jump-full&object_id=104509&local_base=GEN01-MCG02 on 01/03/2013.
- Nummenmaa, L., Hirvonen, J., Parkkola, R., & Hietanen, J. (2008). Is emotional contagion special? An fMRI study on neural systems for affective and cognitive empathy. *NeuroImage*, 43, 571-580.
- O'Brien, E., Konrath, S., Grühn, D., & Hagen, L. (2013). Empathic concern and perspective taking: Linear and quadratic effects of age across the adult life span. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 68, 168-175.
- O'Connell, S. (1995). Empathy in chimpanzees: Evidence for theory of mind? *Primates*, 36, 397-410.
- O'Leary-Kelly, A., Martocchio, J., & Frink, D. (1994). A review of the influence of group goals on group performance. *The Academy of Management Journal*, 37, 1285-1301.
- O'Reilly, C., Caldwell, D., & Barnett, W. (1989). Work group demography, social integration, and turnover. *Administrative Science Quarterly*, 34, 21-37.
- Oppenheimer, D., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45, 867-872.
- Paolini, S., Hewstone, M., Voci, A., Harwood, J., & Cairns, E. (2006). Intergroup contact and the promotion of intergroup harmony: The influence of intergroup emotions. In R. Brown & D. Capozza (Eds.), *Social identities: Motivational, emotional and cultural influences* (pp.209-238). Hove: Taylor & Francis.
- Paulhus, D. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.
- Paulhus, D., Robins, R., Trzesniewski, K., & Tracy, J. (2004). Two replicable suppressor situations in personality research. *Multivariate Behavioral Research*, 39, 301-326.
- Paulus, C. (2009). *Der Saarbrücker Persönlichkeitsfragebogen SPF (IRI) zur Messung von Empathie: Evaluation der deutschen Version des Interpersonal Reactivity Index*. [The Saarbrueck Personality Questionnaire on empathy: Psychometric evaluation of the German version of the Interpersonal Reactivity Index]. Retrieved from: http://www.uni-saarland.de/fak5/ezw/personal/paulus/empathy/SPF_Artikel.pdf on 01/03/2013.

- Pecukonis, E. (1997). A cognitive/affective empathy training program as a function of ego development in aggressive adolescent females. *Adolescence*, 25, 59-76.
- Penke, L. & Asendorpf, J. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, 95, 1113-1135.
- Perkins, D. & Saris, R. (2001). A "Jigsaw Classroom" technique for undergraduate statistics courses. *Teaching of Psychology*, 28, 111-113.
- Peterson, R. (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. *Journal of Consumer Research*, 28, 450-461.
- Pettigrew, T. & Tropp, L. (2005). Allport's intergroup contact hypothesis: Its history and influence. In J. Dovidio, P. Glick, & L. Budman (Eds.), *The nature of prejudice* (pp. 262-277). Malden, MA: Blackwell.
- Plutchik, R. (1987). Evolutionary bases of empathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 38-46). New York: Cambridge University Press.
- Podsakoff, P., MacKenzie, S., Lee, J., & Podsakoff, N. (2006). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88, 879-903.
- Poletti, M., Enrici, I., & Adenzato, M. (2012). Cognitive and affective theory of mind in neurodegenerative diseases: Neuropsychological, neuroanatomical and neurochemical levels. *Neuroscience & Biobehavioral Reviews*, 36, 2147-2164.
- Povinelli, D. (1996). Chimpanzee theory of mind? The long road to strong inference. In P. Carruthers & P. Smith (Eds.), *Theories of theories of mind* (pp. 293-329). New York: Cambridge University Press.
- Prehn-Kristensen, A., Wiesner, C., Bergmann, T., Wolff, S., Jansen, O., Mehdorn, H., Ferstl, R., et al. (2009). Induction of empathy by the smell of anxiety. *Plos One*, 4, e5987.
- Preissler, S., Dziobek, I., Ritter, K., Heekeren, H., & Roepke, S. (2010). Social cognition in Borderline Personality Disorder: Evidence for disturbed recognition of the emotions, thoughts, and intentions of others. *Frontiers in Behavioral Neuroscience*, 4, 182.
- Preston, S. & de Waal, F. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25, 1-72.
- Raafat, R., Chater, N., & Frith, C. (2009). Herding in humans. *Trends in Cognitive Sciences*, 13, 420-428.
- Raff, M., Johnson, A., & Walter, P. (2008) Painful publishing. *Science*, 321, 36.
- Rankin, K., Kramer, J., & Miller, B. (2005). Patterns of cognitive and emotional empathy in frontotemporal lobar degeneration. *Cognitive and Behavioral Neurology*, 18, 28-36.

- Rapisarda, B. (2002). The impact of emotional intelligence on work team cohesiveness and performance. *International Journal of Organizational Analysis*, 10, 363-379.
- Redmond, M. (1989). The functions of empathy (decentering) in human relations. *Human Relations*, 42, 593-605.
- Reiner, R. (1989). *When Harry met Sally* [Motion picture]. Los Angeles: Metro Goldwyn Mayer.
- Repacholi, B. & Gopnik, A. (1997). Early reasoning about desires. Evidence from 14- and 18-month-olds. *Developmental Psychology*, 33, 12-21.
- Repacholi, B. & Slaughter, V. (2003). *Individual differences in theory of mind: Implications for typical and atypical development*. Hove, UK: Psychology Press.
- Rimé, B. (2007). The social sharing of emotion as an interface between individual and collective processes in the construction of emotional climates. *Journal of Social Issues*, 63, 307-322.
- Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review*, 1, 60-85.
- Ritter, K., Dziobek, I., Preißler, S., Rüter, A., Vater, A., Fydrich, T., Lammers, C., et al. (2011). Lack of empathy in patients with narcissistic personality disorder. *Psychiatry Research*, 2011, 241-247.
- Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169-192.
- Roark, A. & Sharah, H. (1989). Factors related to group cohesiveness. *Small Group Behavior*, 20, 62-69.
- Roberts, N., Tsai, J., & Coan, J. (2007). Emotion elicitation using dyadic interaction tasks. In: J. Coan & J. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 106-123). New York: Oxford University Press.
- Romero, T., Castellanos, M., & de Waal, F. (2010). Consolation as possible expression of sympathetic concern among chimpanzees. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 12110-12115.
- Rottenberg, J., Ray, R., & Gross, J. (2007). Emotion elicitation using films. In: J. Coan & J. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 9-28). New York: Oxford University Press.
- Ruby, P. & Decety, J. (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *Journal of Cognitive Neuroscience*, 16, 988-999.
- Ruscio, J., Whitney, D., & Amabile, T. (1998). Looking inside the fishbowl of creativity: Verbal and behavioral predictors of creative performance. *Creativity Research Journal*, 11, 243-263.
- Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161-1178.

- Russell, J., Lewicka, M., & Niit, T. (1989). A cross-cultural study of a circumplex model of affect. *Journal of Personality and Social Psychology*, 57, 848-856.
- Sakia, R. (1992). The Box-Cox transformation technique: A review. *The Statistician*, 41, 169-178.
- Santiesteban, I., White, S., Cook, J., Gilbert, S., Heyes, C., & Bird, G. (2011). Training social cognition: From imitation to theory of mind. *Cognition*, 122, 228-235.
- Saulsman, L. & Page, A. (2004). The five-factor model and personality disorder empirical literature: A meta-analytic review. *Clinical Psychology Review*, 23, 1055-1085.
- Savage, K., Dziobek, I., Teague, E., & Borod, J. (submitted). A new measure of empathy: Psychometric characteristics of the Cognitive and Emotional Empathy Questionnaire. *Personality and Individual Differences*.
- Schafer, J. & Graham, J. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147-177.
- Scheffé, H. (1959). *Analysis of variance*. New York: Wiley.
- Scheithauer, H., Hayer, T., Petermann, F., & Jugert, G. (2006). Physical, verbal, and relational forms of bullying among German students: Age trends, gender differences, and correlates. *Aggressive Behavior*, 32, 261-275.
- Schermuly, C., Schröder, T., Nachtwei, J. & Scholl, W. (2010). Das Instrument zur Kodierung von Diskussionen (IKD). [The Discussion Coding System (DCS)]. *Zeitschrift für Arbeits- und Organisationspsychologie*, 54, 149-170.
- Schermuly, C. & Scholl, W. (2012). The Discussion Coding System (DCS): A new instrument for analyzing communication processes. *Communication Methods and Measures*, 6, 12-40.
- Schieman, S. & van Gundy, K. (2000). The personal and social links between age and self-reported empathy. *Social Psychology Quarterly*, 63, 152-174.
- Schnell, K., Bluschke, S., Konradt, B., & Walter, H. (2011). Functional relations of empathy and mentalizing: An fMRI study on the neural basis of cognitive empathy. *NeuroImage*, 54, 1743-1754.
- Scholl, W. (2013). The socio-emotional basis of human interaction and communication: How we construct our social world. *Social Science Information*, 52, 3-33.
- Schulz, J., Fischbacher, U., Thöni, C., & Utikal, V. (2014). Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology*, 41, 77-87.
- Scott, J. (1988). Social network analysis. *Sociology*, 22, 109-127.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, 10, 70-76.

- Sebastian, C., Fontaine, N., Bird, G., Blakemore, S., Brito, S., McCrory, E., & Viding, E. (2012). Neural processing associated with cognitive and affective theory of mind in adolescents and adults. *Social Cognitive and Affective Neuroscience*, 7, 53-63.
- Seger, C., Smith, E., & Mackie, D. (2009). Subtle activation of a social categorization triggers group-level emotions. *Journal of Experimental Social Psychology*, 45, 460-467.
- Shamay-Tsoory, S., Shur, S., Barcai-Goodman, L., Medlovich, S., Harari, H., & Levkovitz, Y. (2007). Dissociation of cognitive from affective components of theory of mind in schizophrenia. *Psychiatry Research*, 149, 11-23.
- Shamay-Tsoory, S., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, 132, 617-627.
- Sharman, S., Garry, M., & Hunt, M. (2005). Using source cues and familiarity cues to resist imagination inflation. *Acta Psychologica*, 120, 227-242.
- Sherman, K. (2008). *Classroom-based empathy training: An evaluation of program effect in an elementary school*. Ann Arbor: ProQuest.
- Sherman, S., Rose, J., Koch, K., Presson, C., & Chassin, L. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. *Journal of Social and Clinical Psychology*, 22, 13-39.
- Simmons, J., Nelson, L., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 20, 1-8.
- Simon, B., Stürmer, S., & Steffens, K. (2000). Helping individuals or group members? The role of individual and collective identification in AIDS volunteerism. *Personality and Social Psychology Bulletin*, 26, 497-506.
- Simoneau, Y. (2000). *Nuremberg* [Motion picture]. Toronto: Alliance Atlantis Communications.
- Simpson, J., Ickes, W., & Blackstone, T. (1995). When the head protects the heart: Empathic accuracy in dating relationships. *Journal of Personality and Social Psychology*, 69, 629-641.
- Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience and Biobehavioral Reviews*, 30, 855-863.
- Singer, T. & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences*, 1156, 81-96.
- Smith, A. (2006). Cognitive empathy and emotional empathy in human behavior and evolution. *The Psychological Record*, 56, 3-21.

- Smith, A. (2009). The empathy imbalance hypothesis of autism: A theoretical approach to cognitive and emotional empathy in autistic development. *The Psychological Record*, 59, 489-510.
- Smith, E., Seger, C., & Mackie, D. (2007). Can emotions be truly group level? Evidence regarding four conceptual criteria. *Journal of Personality and Social Psychology*, 93, 431-446.
- Sodian, B. & Kristen, S. (2010). Theory of mind. In B. Glatzeder, V. Goel, & A. Müller (Eds.), *Towards a theory of thinking* (pp.189-201). Berlin, Heidelberg: Springer.
- Spencer, S., Steele, C., & Quinn, D. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4-28.
- Spielberg, S. (1993). *Schindler's list* [Motion picture]. Los Angeles: Universal Pictures.
- Steiger, J. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 120, 3-24.
- Stinson, L. & Ickes, W. (1992). Empathic accuracy in the interactions of male friends versus male strangers. *Journal of Personality and Social Psychology*, 62, 787-797.
- Stokes, J. (1983). Components of group cohesion: Intermember attraction, instrumental value, and risk taking. *Small Group Research*, 14, 163-173.
- Stöber, J. (1999). Die soziale Erwünschtheits-Skala-17 (SES-17): Entwicklung und erste Befunde zu Reliabilität und Validität. [The Social Desirability Scale-17 (SES-17): Development and first results on reliability and validity]. *Diagnostica*, 45, 173-177.
- Stöber, J. (2001). The Social Desirability Scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment*, 17, 222-232.
- Stueber, K. (2008). Empathy. In: E. Zalda (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from <http://plato.stanford.edu/archives/fall2008/entries/empathy> on 30/08/2012.
- Stürmer, S., Snyder, M., Kropp, A., & Siem, B. (2006). Empathy-motivated helping: The moderating role of group membership. *Personality and Social Psychology Bulletin*, 32, 943-956.
- Suddendorf, T. & Whiten, A. (2001). Mental evolution and development: Evidence for secondary representation in children, great apes, and other animals. *Psychological Bulletin*, 127, 629-650.
- Sze, J., Gyurak, A., Goodkind, M., & Levenson, R. (2012). Greater emotional empathy and prosocial behavior in late life. *Emotion*, 12, 1129-1140.
- Tabachnik, B. & Fidell, L. (2007). *Using multivariate statistics*. Boston: Allyn & Bacon.
- Thibault, P., Bourgeois, P., & Hess, U. (2006). The effect of group-identification on emotion recognition: The case of cats and basketball players. *Journal of Experimental Social Psychology*, 42, 676-683.
- Thompson, E. (2001). Empathy and consciousness. *Journal of Consciousness Studies*, 8, 1-32.
- Toussaint, L. & Webb, J. (2005). Gender differences in the relationship between empathy and forgiveness. *The Journal of Social Psychology*, 145, 673-685.

- Trevarthen, C. & Aitken, K. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry*, 42, 3-48.
- Tuckman, B. (1965). Developmental sequence in small groups. *Psychological Bulletin*, 63, 384-399.
- Tuckman, B. & Jensen, M. (1977). Stages of small group development revisited. *Group Organization Management*, 2, 419-427.
- Ungerer, J., Dolby, R., Waters, B., Barnett, B., Kelk, N., & Lewin, V. (1990). The early development of empathy: Self-regulation and individual differences in the first year. *Motivation and Emotion*, 14, 93-106.
- van der Zee, K., Thijs, M., & Schakel, L. (2002). The relationship of emotional intelligence with academic intelligence and the Big Five. *European Journal of Personality*, 16, 103-125.
- van Dick, R., van Knippenberg, D., Hägele, S., Guillaume, Y., & Brodbeck, F. (2008). Group diversity and group identification: The moderating role of diversity beliefs. *Human Relations*, 61, 1463-1492.
- van Knippenberg, D., de Dreu, C., & Homan, A. (2004). Work group diversity and group performance: An integrative model and research agenda: Theoretical models and conceptual analyses. *Journal of Applied Psychology*, 89, 1008-1022.
- van Mierlo, H. & Kleingeld, A. (2010). Goals, strategies, and group performance: Some limits of goal setting in groups. *Small Group Research*, 41, 524-555.
- Wacker, R. (2011). *Empathische Kommunikation und sozioemotionale Prozesse in Gruppen*. [Empathic communication and socioemotional processes in groups]. Unpublished master's thesis, Ruhr Universität Bochum, Germany.
- Walter, H., Schnell, K., Erk, S., Arnold, C., Kirsch, P., Esslinger, C., Mier, D., Schmittgen, M., et al. (2011). Effects of a genome-wide supported psychosis risk variant on neural activation during a theory-of-mind task. *Molecular Psychiatry*, 16, 462-470.
- Walter, H. (2012). The social cognitive neuroscience of empathy: Concepts, circuits, and genes. *Emotion Review*, 4, 9-17.
- Wasserman, S. (1994). *Social network analysis: Methods and applications*. New York: Cambridge University Press.
- Wastell, C., Cairns, D., & Haywood, H. (2009). Empathy training, sex offenders and re-offending. *Journal of Sexual Aggression*, 15, 149-159.
- Watson, P. & Morris, R. (1991). Narcissism, empathy and social desirability. *Personality and Individual Differences*, 12, 575-579.
- Webber, S. & Donahue, L. (2001). Impact of highly and less job-related diversity on work group cohesion and performance: A meta-analysis. *Journal of Management*, 27, 141-162.

- Wegge, J. (2000). Participation in group goal setting: Some novel findings and a comprehensive model as a new ending to an old story. *Applied Psychology: An International Review*, 49, 498-516.
- Wegge, J. & Hasslam, S. (2005). Improving work motivation and performance in brainstorming groups: The effects of three group goal-setting strategies. *European Journal of Work and Organizational Psychology*, 14, 400-430.
- Westermann, R., Spies, K., Stahl, G., & Hesse, F. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, 26, 557-580.
- Wheless, L., Wheless, V., & Dickson-Markman, F. (1982). The relations among social and task perceptions in small groups. *Small Group Behavior*, 13, 373-384.
- Wicker, A. (1969). Attitude versus action: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, 25, 41-78.
- Widmeyer, N. & Ducharme, K. (1997). Team building through team goal setting. *Journal of Applied Sport Psychology*, 9, 97-113.
- Wild, B., Erb, M., & Bartels, M. (2001). Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: Quality, quantity, time course and gender differences. *Psychiatry Research*, 102, 109-124.
- Wilkinson, M. (1992). How do we understand empathy systemically? *Journal of Family Therapy*, 14, 193-205.
- Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128.
- Wispé, L. (1986). The distinction between sympathy and empathy: To call forth a concept, a word is needed. *Journal of Personality and Social Psychology*, 50, 314-321.
- Wood, S., Wood, E., & Boyd, S. (2010). *Mastering the world of psychology*. Boston: Allyn & Bacon.
- Wright, R., Contrada, R., & Patane, M. (1986). Task difficulty, cardiovascular response, and the magnitude of goal valence. *Journal of Personality and Social Psychology*, 51, 837-843.
- Yabar, Y. & Hess, U. (2007). Display of empathy and perception of out-group members. *New Zealand Journal of Psychology*, 36, 42-49.
- Yeo, G. & Neal, A. (2004). A multilevel analysis of effort, practice, and performance: Effects of ability, conscientiousness, and goal orientation. *Journal of Applied Psychology*, 89, 231-247.
- Zaki, J., Bolger, N., & Ochsner, K. (2008). It takes two: The interpersonal nature of empathic accuracy. *Psychological Science*, 19, 399-404.

8. Appendix

8.1 Operationalization of Empathic Communication (Wacker, 2011)

Table A.1

Operationalization of Empathic Statements

Dimension	Level 1: Repetition	Level 2: Elaboration / Evaluation	Level 3: Explicit Reaction
Cognitive Empathy	<ul style="list-style-type: none"> - citations - paraphrasing - summarizing 	<ul style="list-style-type: none"> - interpretations - conclusions - presumptions - representative cognitions on behalf of the other <ul style="list-style-type: none"> ○ associations ○ reminiscence ○ verbalizations ○ questions 	explicit statements implying that the speaker understands the emotional states of the other person, adopts the other person's perspective, and comprehends the other person's point of view
Emotional Empathy	nonverbal reactions to the other person's emotional states	<ul style="list-style-type: none"> - acceptance - legitimization - validation - confirmation - signs of approval - appreciation 	explicit statements implying that the speaker isovalently shares the emotions of the other person

Note. For more details and references, see Wacker, 2011.

Table A.2

Operationalization of Empathic Feedback

Dimension	
Nonverbal	<ul style="list-style-type: none"> - nod of the head - smiling / laughing - facial expressions and gestures accompanying, emphasizing, or complementing the other person's messages, e.g. gesticulations, shrug of the shoulders, rolling the eyes, etc.
Paraverbal	<ul style="list-style-type: none"> - sounds as "hm", "aha", etc.
Verbal	<ul style="list-style-type: none"> - initial feedback words, e.g. "sure!", "oh yes!" - short validating expressions, e.g. "oh!", "nice!", "what?!", etc.

Note. For more details and references, see Wacker, 2011.

8.2 Detailed Associations Between Empathy and Control Variables

Table A.3

Bivariate Correlations Between Social Desirability, Big Five Personality Traits, and Empathy in the First Experimental Session

Measures	Big Five Personality Traits					
	SDS	N	E	O	C	A
<i>IRI</i>						
Cognitive Empathy	.08	.08	.08	.18 ⁺	.28**	.26**
- Perspective Taking	.19 ⁺	.01	-.03	.11	.30**	.32**
- Fantasy	-.05	.01	.21*	.07	.17	.08
Emotional Empathy	-.11	.36***	.15	.16	.07	.15
- Empathic Concern	.11	.11	.20 ⁺	.17	.18 ⁺	.30**
- Personal Distress	-.23*	.35**	.06	.05	-.04	-.03
<i>CEEQ-GROUP</i>						
Cognitive Empathy	.04	.06	-.01	.17	.14	.01
- Perspective Taking	.14	.03	-.06	.18 ⁺	.21*	.16
- Mental Perception	-.05	.06	.03	.12	.05	-.12
Emotional Empathy	-.01	.19 ⁺	.01	.03	.15	.19 ⁺
- Empathic Concern	-.04	.20 ⁺	-.03	.10	.15	.22*
- Mirroring	.03	.12	.05	-.05	.12	.12
<i>MET</i>						
Cognitive Empathy	-.32**	.18 ⁺	-.28**	.30**	-.19 ⁺	.02
Emotional Empathy	-.11	.05	.14	.22*	.17	.31**
<i>Public Goods Game</i>						
Given Amount in €	-.09	-.14	.03	.12	.12	.05
<i>DCS</i>						
Emp. Statements	-.05	-.02	-.04	.22*	-.17	-.01
Empathic Feedback	-.06	.04	-.07	.08	.07	.17

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

N=Neuroticism, E=Extraversion, O=Openness, C=Conscientiousness, A=Agreeableness.

Table A.4

Bivariate Correlations Between Social Desirability, Big Five Personality Traits, and Empathy in the Fourth Experimental Session

Measures	SDS		Big Five Personality Traits				
		N	E	O	C	A	
<i>IRI</i>							
Cognitive Empathy	.16	.12	.05	.30 **	.19 ⁺	.27 **	
- Perspective Taking	.23 *	.04	.00	.23 *	.23 *	.35 **	
- Fantasy	.04	.15	.08	.26 *	.11	.12	
Emotional Empathy	-.12	.47 ***	.05	.24 *	.01	.18 ⁺	
- Empathic Concern	.05	.26 *	.03	.24 *	.13	.26 *	
- Personal Distress	-.23 *	.48 ***	.05	.15	-.11	.05	
<i>CEEQ-GROUP</i>							
Cognitive Empathy	.22 *	.05	.11	.21 *	.25 *	.19 ⁺	
- Perspective Taking	.21 *	.09	.06	.20 ⁺	.24 *	.22 *	
- Mental Perception	.16	.00	.12	.18 ⁺	.18 ⁺	.13	
Emotional Empathy	.06	.10	.18	-.01	.24 *	.23 *	
- Empathic Concern	.05	.06	.09	.09	.25 *	.30 **	
- Mirroring	.06	.12	.24 *	-.11	.18 ⁺	.11	
<i>MET</i>							
Cognitive Empathy	-.21 ⁺	.17	-.25 *	.26 *	-.18 ⁺	.03	
Emotional Empathy	-.13	.08	.13	.17	.24 *	.26 *	
<i>Public Goods Game</i>							
Given Amount in €	.03	-.30 **	.16	-.08	.24 *	.02	
<i>DCS</i>							
Emp. Statements	-.04	.05	.04	.36 ***	.10	.14	
Empathic Feedback	.22 *	-.06	.08	.09	.15	.30 **	

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

N=Neuroticism, E=Extraversion, O=Openness, C= Conscientiousness, A=Agreeableness.

Table A.5

Bivariate Correlations Between Social Desirability, Big Five Personality Traits, and ΔM in Empathy Measures Between the First and the Fourth Experimental Session

Measures	SDS		Big Five Personality Traits				
		N	E	O	C	A	
<i>IRI</i>							
Cognitive Empathy	.13	.08	-.03	.21 *	-.11	.03	
- Perspective Taking	.03	.04	.04	.15	-.15	-.01	
- Fantasy	.12	.14	-.14	.21 *	-.06	.03	
Emotional Empathy	.00	.19 ⁺	-.12	.11	-.06	.03	
- Empathic Concern	-.07	.23 *	-.24 *	.14	-.03	-.03	
- Personal Distress	.04	.14	-.01	.08	-.06	.09	
<i>CEEQ-GROUP</i>							
Cognitive Empathy	.23 *	.00	.16	.05	.11	.24 *	
- Perspective Taking	.07	.07	.14	-.01	-.02	.04	
- Mental Perception	.28 **	-.05	.13	.07	.17	.31 **	
Emotional Empathy	.08	-.13	.24 *	-.07	.10	.04	
- Empathic Concern	.11	-.20 ⁺	.15	-.04	.10	.08	
- Mirroring	.03	-.02	.23 *	-.07	.06	-.02	
<i>MET</i>							
Cognitive Empathy	.13	.04	-.02	.00	-.04	.02	
Emotional Empathy	-.05	.08	.03	-.02	.18 ⁺	.05	
<i>Public Goods Game</i>							
Given Amount in €	.10	-.21 *	.13	-.16	.14	-.04	
<i>DCS</i>							
Emp. Statements	.01	.05	.07	.10	.21 *	.12	
Empathic Feedback	.27 **	-.12	.15	.00	.08	.14	

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

N=Neuroticism, E=Extraversion, O=Openness, C=Conscientiousness, A=Agreeableness.

Table A.6

Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation, and Empathy (First Session)

Measures	GC	GI	DO	AF	PD	DB	GR
<i>IRI</i>							
Cognitive Empathy	.12	.33 **	.02	.03	-.09	.09	.03
- Perspective Taking	.14	.25 *	-.08	.01	-.15	.02	.13
- Fantasy	.06	.24 *	.05	.00	-.02	.09	-.02
Emotional Empathy	-.07	.17	-.06	.04	.10	-.06	-.02
- Empathic Concern	-.04	.16	-.06	-.06	.23 *	.05	.20 *
- Personal Distress	-.05	.10	-.04	.10	-.08	-.12	-.17
<i>CEEQ-GROUP</i>							
Cognitive Empathy	.26 *	.19 ⁺	-.08	.10	.00	.06	.03
- Perspective Taking	.26 *	.19 ⁺	-.14	-.01	-.06	.06	.06
- Mental Perception	.20 ⁺	.15	-.03	.17	.05	.05	-.01
Emotional Empathy	.07	.37 ***	-.07	.15	.26*	.02	-.02
- Empathic Concern	.03	.39 ***	-.05	.16	.26 *	.05	-.05
- Mirroring	.09	.28 **	-.06	.10	.19 ⁺	-.02	.01
<i>MET</i>							
Cognitive Empathy	-.24 *	-.10	.41 ***	.24 *	.14	.00	-.22 *
Emotional Empathy	.04	.30 **	-.04	-.18 ⁺	.00	.03	.05
<i>Public Goods Game</i>							
Given Amount in €	.06	-.18 ⁺	-.12	-.27 **	.06	-.04	.07
<i>DCS</i>							
Emp. Statements	.09	.19 ⁺	.14	.09	.00	-.02	-.17 ⁺
Empathic Feedback	.10	.05	-.05	.18 ⁺	-.15	-.08	.15

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=98$.

GC=Group Climate, GI=Group Identification, DO=Dominance, AF= Affiliation, PD= Perceived Diversity, DB=Diversity Beliefs, GR=Goal Representation.

Table A.7

Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation, and Empathy (Fourth Session)

Measures	GC	GI	DO	AF	PD	DB	GR
<i>IRI</i>							
Cognitive Empathy	.17	.29 **	.04	.20 ⁺	-.05	.08	-.13
- Perspective Taking	.18 ⁺	.26 *	-.04	.05	-.03	.11	-.06
- Fantasy	.10	.22 *	.10	.28 **	-.05	.01	-.17
Emotional Empathy	.11	.15	.06	.23 *	.13	-.05	-.09
- Empathic Concern	.17	.08	.07	.26 *	.12	-.01	-.03
- Personal Distress	.01	.15	-.01	.11	.08	-.11	-.11
<i>CEEQ-GROUP</i>							
Cognitive Empathy	.25 *	.30 **	.05	.21 *	.08	.10	.11
- Perspective Taking	.14	.16	.10	.17	.07	.20 ⁺	-.08
- Mental Perception	.27 **	.33 **	.01	.18 ⁺	.06	.01	.21
Emotional Empathy	.19 ⁺	.44 ***	.00	.29 **	.16	-.15	.00
- Empathic Concern	.19 ⁺	.35 **	-.06	.31 **	.11	-.09	.10
- Mirroring	.16	.44 ***	.05	.23 *	.18 ⁺	-.19 ⁺	-.10
<i>MET</i>							
Cognitive Empathy	-.10	-.17	.09	.17 ⁺	.10	.04	-.17
Emotional Empathy	.12	.26 *	.07	.31 **	.02	-.16	.16
<i>Public Goods Game</i>							
Given Amount in €	.32 **	.26 *	.19 ⁺	.14	-.10	.01	.24 *
<i>DCS</i>							
Emp. Statements	.06	-.01	.41 ***	.49 ***	-.13	-.02	-.17
Empathic Feedback	.02	.23 *	-.01	.02	-.04	-.05	.07

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

GC=Group Climate, GI=Group Identification, DO=Dominance, AF= Affiliation, PD= Perceived Diversity, DB=Diversity Beliefs, GR=Goal Representation.

Table A.8

Bivariate Correlations Between ΔM in Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation, and ΔM in Empathy

Measures	GC	GI	DO	AF	PD	DB	GR
<i>IRI</i>							
Cognitive Empathy	.05	.07	-.08	-.09	.20 ⁺	.01	.00
- Perspective Taking	-.03	.08	-.05	-.02	.16	-.03	-.03
- Fantasy	-.02	-.03	-.16	-.11	.17	.01	.06
Emotional Empathy	-.16	-.26 [*]	-.19 ⁺	.04	.01	-.30 ^{**}	.12
- Empathic Concern	-.05	-.18 ⁺	-.13	.11	.07	-.13	.20 ⁺
- Personal Distress	-.18 ⁺	-.23 [*]	-.20 ⁺	.01	-.04	-.34 ^{**}	-.01
<i>CEEQ-GROUP</i>							
Cognitive Empathy	.13	.09	.08	.26 [*]	-.08	.18 ⁺	-.04
- Perspective Taking	.00	.03	.11	.22 [*]	-.22 [*]	.26 [*]	-.14
- Mental Perception	.19 ⁺	.11	.03	.20 ⁺	.06	.05	.05
Emotional Empathy	.10	.18 ⁺	-.08	.21 [*]	.04	-.03	-.22 [*]
- Empathic Concern	.08	.20 ⁺	-.10	.13	-.08	.02	-.23 [*]
- Mirroring	.08	.10	-.03	.20 ⁺	.13	-.06	-.13
<i>MET</i>							
Cognitive Empathy	.01	.10	-.06	-.16	.09	-.14	.12
Emotional Empathy	.16	.13	.16	.14	-.03	.06	-.10
<i>Public Goods Game</i>							
Given Amount in €	.18 ⁺	.03	.05	.05	.19 ⁺	-.16	.16
<i>DCS</i>							
Emp. Statements	.13	.09	-.03	.29 ^{**}	-.25 [*]	-.03	-.06
Empathic Feedback	.04	.10	.04	.04	-.04	.00	.04

Note. ⁺ $p < .10$; ^{*} $p < .05$; ^{**} $p < .01$; ^{***} $p < .001$ (two-tailed). $N=91$.

GC=Group Climate, GI=Group Identification, DO=Dominance, AF= Affiliation, PD= Perceived Diversity, DB=Diversity Beliefs, GR=Goal Representation.

Table A.9

Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation in First session and Empathy in Fourth session

Measures	GC	GI	DO	AF	PD	DB	GR
<i>IRI</i>							
Cognitive Empathy	.08	.33 **	.08	.13	-.14	.21 *	-.01
- Perspective Taking	.12	.26 *	-.05	.11	-.13	.18 ⁺	.08
- Fantasy	.00	.30 **	.16	.12	-.10	.15	-.09
Emotional Empathy	-.06	.24 *	-.02	.00	.20 ⁺	.06	-.08
- Empathic Concern	-.05	.16	-.07	-.10	.21 *	.17	.07
- Personal Distress	-.07	.20 ⁺	.01	.09	.10	-.08	-.17
<i>CEEQ-GROUP</i>							
Cognitive Empathy	.18 ⁺	.23 *	-.08	-.07	.03	.05	.11
- Perspective Taking	.21 *	.20 ⁺	-.03	-.10	.07	.04	-.02
- Mental Perception	.12	.20 ⁺	-.09	-.04	-.01	.05	.16
Emotional Empathy	.08	.43 ***	-.05	.05	.13	.06	.16
- Empathic Concern	.03	.32 **	-.10	.08	.15	.10	.22 *
- Mirroring	.11	.46 ***	.01	-.01	.09	.00	.05
<i>MET</i>							
Cognitive Empathy	-.27 **	-.19 ⁺	.37 ***	.26 *	.02	.10	-.20 ⁺
Emotional Empathy	.04	.22 *	-.06	-.24 *	.01	-.02	.30 **
<i>Public Goods Game</i>							
Given Amount in €	.14	.03	-.04	-.25 *	-.17 ⁺	.06	.13
<i>DCS</i>							
Emp. Statements	.02	.06	.46 ***	.03	-.04	.00	-.20 ⁺
Empathic Feedback	.13	.19 ⁺	-.12	.05	-.02	-.02	.03

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

GC=Group Climate, GI=Group Identification, DO=Dominance, AF= Affiliation, PD= Perceived Diversity, DB=Diversity Beliefs, GR=Goal Representation.

Table A.10

Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation in Fourth session and Empathy in First session

Measures	GC	GI	DO	AF	PD	DB	GR
<i>IRI</i>							
Cognitive Empathy	.19 ⁺	.26 [*]	.06	.16	-.12	.01	-.05
- Perspective Taking	.26 [*]	.27 [*]	-.01	-.05	-.11	.06	.06
- Fantasy	.20 ⁺	.23 [*]	.16	.24 [*]	-.13	-.06	-.12
Emotional Empathy	.28 ^{**}	.27 ^{**}	.16	.23 [*]	.00	-.07	-.11
- Empathic Concern	.28 ^{**}	.20 ⁺	.17 ⁺	.22 [*]	.10	-.11	-.01
- Personal Distress	.20 ⁺	.25 [*]	.09	.10	-.11	-.01	-.12
<i>CEEQ-GROUP</i>							
Cognitive Empathy	.22 [*]	.21 [*]	-.02	.07	.08	.08	.13
- Perspective Taking	.21 [*]	.17	-.07	-.04	.04	.07	.15
- Mental Perception	.16	.19 ⁺	.02	.14	.10	.08	.08
Emotional Empathy	.17	.27 ^{**}	.02	.14	.29 ^{**}	-.10	.01
- Empathic Concern	.18 ⁺	.26 [*]	.02	.18 ⁺	.31 ^{**}	-.10	.05
- Mirroring	.13	.24 [*]	.02	.08	.21 [*]	-.08	-.02
<i>MET</i>							
Cognitive Empathy	-.13	-.20 ⁺	.14	.25 [*]	.17	.03	-.22 [*]
Emotional Empathy	.01	.21 [*]	.00	.32 ^{**}	-.02	-.15	-.04
<i>Public Goods Game</i>							
Given Amount in €	.09	.08	.10	.14	.01	.13	.02
<i>DCS</i>							
Emp. Statements	-.07	-.04	.22 [*]	-.03	.11	-.02	-.04
Empathic Feedback	-.01	.05	.01	.06	-.18 ⁺	-.15	.15

Note. ⁺ $p < .10$; ^{*} $p < .05$; ^{**} $p < .01$; ^{***} $p < .001$ (two-tailed). $N=91$.

GC=Group Climate, GI=Group Identification, DO=Dominance, AF= Affiliation, PD= Perceived Diversity, DB=Diversity Beliefs, GR=Goal Representation.

Table A.11

Bivariate Correlations Between ΔM in Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation, and Empathy in the First session

Measures	GC	GI	DO	AF	PD	DB	GR
<i>IRI</i>							
Cognitive Empathy	.12	-.04	.03	.10	.02	-.15	-.14
- Perspective Taking	.15	.02	.03	-.05	.06	-.05	-.12
- Fantasy	.15	.00	.14	.18 ⁺	-.05	-.14	-.14
Emotional Empathy	.28 ^{**}	.06	.25 [*]	.12	-.12	.09	-.10
- Empathic Concern	.26 [*]	-.01	.25 [*]	.20 ⁺	-.16	-.13	-.26 [*]
- Personal Distress	.20 ⁺	.12	.15	.00	-.03	.23 [*]	.07
<i>CEEQ-GROUP</i>							
Cognitive Empathy	.06	.02	.05	-.01	.15	-.08	.05
- Perspective Taking	.06	-.02	.02	-.02	.17	-.07	.06
- Mental Perception	.04	.05	.05	-.02	.09	-.06	.02
Emotional Empathy	.09	-.07	.08	.00	.01	-.19 ⁺	.01
- Empathic Concern	.13	-.08	.07	.02	.02	-.20 ⁺	.06
- Mirroring	.04	-.03	.07	-.02	-.01	-.15	-.05
<i>MET</i>							
Cognitive Empathy	.07	-.04	-.16	.02	.07	.02	-.04
Emotional Empathy	.00	-.06	.01	.35 ^{***}	.09	-.21 ⁺	-.12
<i>Public Goods Game</i>							
Given Amount in €	.04	.25 [*]	.19 ⁺	.30 ^{**}	-.08	.12	-.06
<i>DCS</i>							
Emp. Statements	-.11	.17	.14	-.08	.19 ⁺	.02	.11
Empathic Feedback	-.08	-.02	.03	-.09	-.01	-.04	.01

Note. ⁺ $p < .10$; ^{*} $p < .05$; ^{**} $p < .01$; ^{***} $p < .001$ (two-tailed). $N=91$.

GC=Group Climate, GI=Group Identification, DO=Dominance, AF= Affiliation, PD= Perceived Diversity, DB=Diversity Beliefs, GR=Goal Representation.

Table A.12

Bivariate Correlations Between ΔM in Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation, and Empathy in the Fourth session

Measures	GC	GI	DO	AF	PD	DB	GR
<i>IRI</i>							
Cognitive Empathy	.14	.01	-.04	.04	.16	-.15	-.13
- Perspective Taking	.13	.06	-.03	-.05	.17	-.09	-.16
- Fantasy	.11	-.05	-.03	.10	.10	-.16	-.10
Emotional Empathy	.15	-.12	.08	.15	-.09	-.13	-.01
- Empathic Concern	.21 *	-.11	.13	.23 *	-.10	-.20 ⁺	-.12
- Personal Distress	.06	-.05	-.02	.01	-.04	-.02	.08
<i>CEEQ-GROUP</i>							
Cognitive Empathy	.15	.08	.11	.18 ⁺	.07	.07	.01
- Perspective Taking	.04	-.01	.13	.18 ⁺	-.01	.16	-.07
- Mental Perception	.19 ⁺	.13	.06	.15	.12	-.02	.05
Emotional Empathy	.17	.07	.02	.16	.04	-.22 *	-.17
- Empathic Concern	.20 ⁺	.08	-.01	.14	-.04	-.20 ⁺	-.15
- Mirroring	.11	.06	.04	.15	.11	-.20 ⁺	-.17
<i>MET</i>							
Cognitive Empathy	.06	.02	-.16	-.06	.11	-.07	.04
Emotional Empathy	.10	.04	.11	.36 ***	.05	-.13	-.17
<i>Public Goods Game</i>							
Given Amount in €	.22 *	.22 *	.22 *	.26 *	.12	-.03	.12
<i>DCS</i>							
Emp. Statements	.04	-.06	.11	.30 **	-.14	-.02	.03
Empathic Feedback	-.04	.07	.06	-.03	-.05	-.04	.05

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

GC=Group Climate, GI=Group Identification, DO=Dominance, AF= Affiliation, PD= Perceived Diversity, DB=Diversity Beliefs, GR=Goal Representation.

Table A.13

Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation in First session and ΔM in Empathy

Measures	GC	GI	DO	AF	PD	DB	GR
<i>IRI</i>							
Cognitive Empathy	-.04	.03	.08	.18 ⁺	-.04	.10	-.12
- Perspective Taking	-.08	-.08	.02	.13	.01	.07	-.16
- Fantasy	-.04	.07	.13	.16	-.02	.09	-.10
Emotional Empathy	.01	.14	.07	-.07	.20 ⁺	.26 [*]	-.08
- Empathic Concern	-.06	.05	.01	-.08	.02	.27 ^{**}	-.22 [*]
- Personal Distress	.04	.14	.06	-.04	.27 ^{**}	.16	.06
<i>CEEQ-GROUP</i>							
Cognitive Empathy	-.11	.05	.03	-.19 ⁺	.05	-.12	-.02
- Perspective Taking	-.11	-.04	.16	-.07	.17 ⁺	-.10	-.16
- Mental Perception	-.07	.10	-.08	-.22 [*]	-.06	-.09	.10
Emotional Empathy	-.02	.07	.02	-.12	-.16	-.01	.17
- Empathic Concern	-.04	-.09	-.04	-.05	-.13	.02	.28 ^{**}
- Mirroring	.00	.18 ⁺	.07	-.14	-.12	-.03	.02
<i>MET</i>							
Cognitive Empathy	-.04	-.14	-.02	.14	-.13	.15	-.07
Emotional Empathy	.00	-.01	-.09	-.12	.11	-.08	.38 ^{***}
<i>Public Goods Game</i>							
Given Amount in €	.09	.17	.04	-.06	-.24 [*]	.04	.09
<i>DCS</i>							
Emp. Statements	-.03	-.06	.24 [*]	-.05	-.03	.03	-.04
Empathic Feedback	.06	.11	-.10	-.13	.13	.11	-.09

Note. ⁺ $p < .10$; ^{*} $p < .05$; ^{**} $p < .01$; ^{***} $p < .001$ (two-tailed). $N=91$.

GC=Group Climate, GI=Group Identification, DO=Dominance, AF= Affiliation, PD= Perceived Diversity, DB=Diversity Beliefs, GR=Goal Representation.

Table A.14

Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation in Fourth session and ΔM in Empathy

Measures	GC	GI	DO	AF	PD	DB	GR
<i>IRI</i>							
Cognitive Empathy	-.01	.08	-.02	.05	.09	.11	-.12
- Perspective Taking	-.12	-.02	-.04	.11	.12	.06	-.19 ⁺
- Fantasy	-.08	.00	-.05	.00	.08	.10	-.05
Emotional Empathy	-.20 ⁺	-.14	-.12	.00	.20 ⁺	.02	.03
- Empathic Concern	-.13	-.15	-.10	.09	.04	.18 ⁺	-.05
- Personal Distress	-.21 [*]	-.10	-.14	-.03	.26 [*]	-.14	.05
<i>CEEQ-GROUP</i>							
Cognitive Empathy	.05	.11	.09	.18 ⁺	-.01	.03	-.06
- Perspective Taking	-.10	-.04	.19 ⁺	.25 [*]	.03	.14	-.28 ^{**}
- Mental Perception	.14	.20 ⁺	-.02	.08	-.03	-.07	.14
Emotional Empathy	.05	.22 [*]	-.04	.20 ⁺	-.16	-.07	-.02
- Empathic Concern	.02	.10	-.11	.14	-.24 [*]	.02	.08
- Mirroring	.05	.24 [*]	.03	.17	-.03	-.11	-.09
<i>MET</i>							
Cognitive Empathy	.01	-.02	-.07	-.10	-.07	.02	.04
Emotional Empathy	.18 ⁺	.15	.09	.08	.07	-.05	.30 ^{**}
<i>Public Goods Game</i>							
Given Amount in €	.24 [*]	.20 ⁺	.09	.03	-.11	-.12	.23 [*]
<i>DCS</i>							
Emp. Statements	.11	.04	.14	.39 ^{***}	-.19 ⁺	.00	-.09
Empathic Feedback	.04	.19 ⁺	-.04	-.07	.12	-.13	-.05

Note. ⁺ $p < .10$; ^{*} $p < .05$; ^{**} $p < .01$; ^{***} $p < .001$ (two-tailed). $N=91$.

GC=Group Climate, GI=Group Identification, DO=Dominance, AF= Affiliation, PD= Perceived Diversity, DB=Diversity Beliefs, GR=Goal Representation.

Table A.15

Associations Between Age, Gender, Education, and Empathy in the First Session

Measures						
	Age		Gender ^b			Education
	Age	Age ^{2 a}	M(SD) ♂	M(SD) ♀	p	
<i>IRI</i>						
Cognitive Empathy	.24 *	-.50	3.35 (.51)	3.50 (.45)	.14	.23 *
- Perspective Taking	.33 **	-1.17	3.50 (.68)	3.57 (.68)	.62	.32 **
- Fantasy	.09	-.69	3.20 (.63)	3.39 (.68)	.22	.06
Emotional Empathy	.06	-2.15 *	2.95 (.41)	3.24 (.40)	.001	.19 ⁺
- Empathic Concern	.13	-1.55 ⁺	3.36 (.53)	3.57 (.49)	.044	.11
- Personal Distress	.01	-1.87 *	2.55 (.61)	2.91 (.61)	.006	.19 ⁺
<i>CEEQ-GROUP</i>						
Cognitive Empathy	.04	.20	3.30 (.45)	3.19 (.57)	.27	.02
- Perspective Taking	.17	-.31	3.27 (.52)	3.30 (.60)	.86	.13
- Mental Perception	-.07	.64	3.33 (.57)	3.10 (.67)	.068	-.07
Emotional Empathy	.29 **	.31	3.14 (.53)	3.24 (.50)	.34	.18 ⁺
- Empathic Concern	.28 **	.48	3.45 (.61)	3.57 (.59)	.35	.13
- Mirroring	.23 *	.10	2.87 (.56)	2.95 (.55)	.46	.18 ⁺
<i>MET</i>						
Cognitive Empathy	.13	1.62 ⁺	.49 (.16)	.52 (.17)	.48	.04
Emotional Empathy	.29 **	.40	4.77 (1.29)	5.02 (1.13)	.26	.08
<i>Public Goods Game</i>						
Given Amount in €	.15	.74	3.16 (1.17)	2.63 (1.20)	.075	.04
<i>DCS</i>						
Emp. Statements	.22 *	1.22	.20 (.22)	.15 (.19)	.32	.04
Empathic Feedback	.10	-.38	.25 (.20)	.51 (.35)	.000	.22 *

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=98$.

^a Non-linear associations between empathy and age were investigated by regressions with age and squared age as predictors, the coefficients here are standardized β .

^b Associations between empathy and gender were investigated by univariate ANOVAs with gender as independent variable and empathy measures as dependent variables.

Table A.16

Associations Between Age, Gender, Education, and Empathy in the Fourth Session

Measures						
	Age		Gender ^b			Education
	Age	Age ^{2 a}	M(SD) ♂	M(SD) ♀	p	
<i>IRI</i>						
Cognitive Empathy	.26 *	.58	3.33 (.58)	3.47 (.49)	.19	.23 *
- Perspective Taking	.34 **	-.18	3.51 (.67)	3.60 (.61)	.46	.26 *
- Fantasy	.10	1.07	3.15 (.72)	3.34 (.62)	.14	.12
Emotional Empathy	.21	-1.40	2.85 (.46)	3.17 (.40)	.000	.24 *
- Empathic Concern	.26 *	-1.08	3.22 (.64)	3.53 (.50)	.009	.15
- Personal Distress	.08	-.81	2.45 (.64)	2.81 (.56)	.003	.23 *
<i>CEEQ-GROUP</i>						
Cognitive Empathy	.12	1.03	3.33 (.50)	3.24 (.53)	.41	-.02
- Perspective Taking	.25 *	.79	3.35 (.51)	3.34 (.51)	.97	.07
- Mental Perception	.01	1.02	3.31 (.66)	3.15 (.70)	.26	-.08
Emotional Empathy	.22 *	.76	3.16 (.59)	3.24 (.47)	.39	.14
- Empathic Concern	.25 *	.79	3.38 (.68)	3.60 (.54)	.078	.15
- Mirroring	.15	.57	2.96 (.60)	2.92 (.52)	.83	.09
<i>MET</i>						
Cognitive Empathy	.12	1.44	.51 (.18)	.53 (.21)	.76	.07
Emotional Empathy	.32 **	-.16	4.65 (1.45)	5.17 (1.50)	.13	.27 *
<i>Public Goods Game</i>						
Given Amount in €	-.02	-.94	3.25 (1.31)	2.67 (1.51)	.096	.02
<i>DCS</i>						
Emp. Statements	-.02	.53	.12 (.18)	.16 (.21)	.34	.03
Empathic Feedback	.13	-.93	.31 (.32)	.47 (.39)	.042	.12

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

^a Non-linear associations between empathy and age were investigated by regressions with age and squared age as predictors, the coefficients here are standardized β .

^b Associations between empathy and gender were investigated by univariate ANOVAs with gender as independent variable and empathy measures as dependent variables.

Table A.17

Associations Between Age, Gender, Education, and ΔM in Empathy Measures

Measures						
	Age		Gender ^b			Education
	Age	Age ^{2 a}	M(SD) ♂	M(SD) ♀	p	
<i>IRI</i>						
Cognitive Empathy	.08	1.25	-.01 (.41)	-.04 (.33)	.71	.05
- Perspective Taking	.01	.69	+.02 (.44)	-.02 (.42)	.66	-.10
- Fantasy	.02	2.24 *	-.03 (.52)	-.02 (.57)	.90	.07
Emotional Empathy	.16	.75	-.10 (.28)	-.06 (.33)	.54	.04
- Empathic Concern	.16	.37	-.13 (.33)	-.02 (.37)	.16	.00
- Personal Distress	.08	1.17	-.11 (.43)	-.09 (.49)	.82	.04
<i>CEEQ-GROUP</i>						
Cognitive Empathy	.08	.57	+.04 (.37)	+.04 (.41)	.99	-.06
- Perspective Taking	-.01	.95	+.09 (.46)	+.03 (.44)	.53	-.13
- Mental Perception	.12	.12	-.02 (.46)	+.04 (.56)	.64	.01
Emotional Empathy	-.11	.03	+.03 (.42)	-.02 (.36)	.60	-.07
- Empathic Concern	-.09	-.46	-.05 (.53)	+.02 (.43)	.50	.01
- Mirroring	-.09	.43	+.10 (.46)	-.05 (.50)	.17	-.12
<i>MET</i>						
Cognitive Empathy	.05	.11	+.03 (.11)	+.02 (.10)	.46	.05
Emotional Empathy	.14	-.05	-.06 (.95)	+.19 (.89)	.20	.32 **
<i>Public Goods Game</i>						
Given Amount in €	-.14	-1.00	+.08 (1.08)	+.04 (1.57)	.89	-.02
<i>DCS</i>						
Emp. Statements	-.15	-.54	-.07 (.27)	+.01 (.24)	.16	.01
Empathic Feedback	.03	-.45	+.07 (.32)	-.04 (.34)	.14	-.13

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=98$.

^a Non-linear associations between empathy and age were investigated by regressions with age and squared age as predictors, the coefficients here are standardized β .

^b Associations between empathy and gender were investigated by univariate ANOVAs with gender as independent variable and empathy measures as dependent variables.

8.3 Correlation Tables for Individual and Systemic Empathy Measures

Table A.18

Partial Correlations Between the IRI Subscales Perspective Taking and Fantasy and the Remaining Empathy Measures

Measures	PT T1	PT T2	PT Δ	F T1	F T2	F Δ
<i>IRI</i>						
Empathic Concern T1	.10	.12	.01	.23 *	.13	-.13
Empathic Concern T2	-.02	.14	.19	.34 **	.35 **	.01
Empathic Concern Δ	-.19	.05	.32 **	.26 *	.40 **	.16
Personal Distress T1	.06	.00	-.06	.08	.13	.08
Personal Distress T2	-.15	-.09	.08	.10	.26 *	.22 ⁺
Personal Distress Δ	-.25 *	-.10	.16	-.01	.14	.20
<i>CEEQ-GROUP</i>						
Perspective Taking T1	.37 **	.59 ***	.23 ⁺	.37 **	.37 **	-.03
Perspective Taking T2	.31 **	.53 ***	.27 *	.32 **	.38 **	.05
Perspective Taking Δ	-.16	-.19 ⁺	-.02	-.16	-.08	.11
Mental State Perception T1	.11	.12	.01	.21 ⁺	.21 ⁺	-.01
Mental State Perception T2	.16	.20 ⁺	.04	.18	.27 *	.11
Mental State Perception Δ	.06	.08	.04	-.03	.08	.14
Empathic Concern T1	.18	.24 *	.10	.27 *	.16	-.14
Empathic Concern T2	.10	.33 **	.27 *	.40 ***	.44 ***	.01
Empathic Concern Δ	-.08	.09	.17	.16	.30 *	.15
Mirroring T1	-.01	.14	.23 *	.17	.24 *	.09
Mirroring T2	.01	.19	.24 *	.31 **	.43 ***	.14
Mirroring Δ	.01	.01	-.05	.11	.14	.03
<i>MET</i>						
Cognitive Empathy T1	-.01	.12	.16	.26 *	.33 **	.08
Cognitive Empathy T2	-.07	.07	.19	.09	.20 ⁺	.13
Cognitive Empathy Δ	-.08	-.02	.10	-.20 ⁺	-.11	.12
Emotional Empathy T1	.19	.15	-.09	.28 *	.21 ⁺	-.10
Emotional Empathy T2	.16	.08	-.14	.24 *	.21 ⁺	-.05
Emotional Empathy Δ	.05	-.01	-.09	.02	.07	.08
<i>Public Goods Game</i>						
Given Amount in € T1	-.06	-.06	.03	.01	-.02	-.02
Given Amount in € T2	-.07	-.05	.05	.12	.09	-.02
Given Amount in € Δ	-.01	.00	.03	.14	.13	-.02
<i>DCS</i>						
Empathic Statements T1	-.08	.16	.35 **	.00	.14	.20 ⁺
Empathic Statements T2	.00	.06	.06	.22 ⁺	.19	-.04
Empathic Statements Δ	.06	-.07	-.22 ⁺	.15	.02	-.19
Empathic Feedback T1	.08	.08	.00	.02	-.01	-.03
Empathic Feedback T2	.13	.09	-.07	-.17	-.21 ⁺	-.08
Empathic Feedback Δ	.09	.02	-.10	-.24 ⁺	-.25 *	-.03

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

T1=First Session, T2=Fourth Session, PT=Perspective Taking, F=Fantasy.

Table A.19

Partial Correlations Between the IRI Subscales Empathic Concern and Personal Distress and the Remaining Empathy Measures

Measures	EC T1	EC T2	EC Δ	PD T1	PD T2	PD Δ
<i>CEEQ-GROUP</i>						
Perspective Taking T1	.09	.12	.08	-.13	-.10	.02
Perspective Taking T2	.14	.20	.12	-.06	.00	.06
Perspective Taking Δ	.05	.07	.06	.10	.15	.06
Mental State Perception T1	.11	.10	-.01	-.05	-.02	.02
Mental State Perception T2	.25 *	.28 *	.10	-.15	-.01	.15
Mental State Perception Δ	.20 ⁺	.22 ⁺	.11	-.13	.02	.17
Empathic Concern T1	.45 ***	.24 *	-.24 ⁺	.14	.15	.03
Empathic Concern T2	.40 ***	.42 ***	.10	.07	.15	.10
Empathic Concern Δ	-.05	.20	.36 **	-.08	-.04	.02
Mirroring T1	.21 ⁺	.10	-.12	.17	.09	-.06
Mirroring T2	.30 *	.32 **	.10	.28 *	.36 **	.16
Mirroring Δ	.06	.20 ⁺	.24 *	.06	.25 *	.24 *
<i>MET</i>						
Cognitive Empathy T1	-.06	.02	.12	-.01	.07	.08
Cognitive Empathy T2	-.07	-.05	.03	.00	.02	.01
Cognitive Empathy Δ	-.02	-.09	-.09	.02	-.05	-.09
Emotional Empathy T1	.40 **	.33 **	-.05	.08	.05	-.03
Emotional Empathy T2	.27 *	.25 *	.00	.11	.20 ⁺	.12
Emotional Empathy Δ	-.10	-.04	.06	.10	.26 *	.21 ⁺
<i>Public Goods Game</i>						
Given Amount in € T1	-.09	-.05	.08	.13	.08	-.04
Given Amount in € T2	.06	.11	.11	-.09	-.12	-.10
Given Amount in € Δ	.13	.14	.07	-.20 ⁺	-.21 ⁺	-.08
<i>DCS</i>						
Empathic Statements T1	.00	.06	.14	.04	-.13	-.24 *
Empathic Statements T2	.13	.21 ⁺	.19	-.15	-.15	-.01
Empathic Statements Δ	.09	.10	.04	-.15	-.02	.18
Empathic Feedback T1	-.11	-.08	.01	.03	-.08	-.14
Empathic Feedback T2	.21 ⁺	.15	-.06	-.14	-.15	-.01
Empathic Feedback Δ	.33 **	.23 ⁺	-.09	-.18	-.11	.11

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

T1=First Session, T2=Fourth Session, EC=Empathic Concern, PD=Personal Distress.

Table A.20

Partial Correlations Between the CEEQ-GROUP Subscales Perspective Taking and Mental State Perception and the Remaining Empathy Measures

Measures	PT T1	PT T2	PT Δ	MP T1	MP T2	MP Δ
<i>CEEQ-GROUP</i>						
Empathic Concern T1	.39 **	.51 ***	.03	.23 ⁺	.31 *	.16
Empathic Concern T2	.43 ***	.47 ***	-.16	.11	.37 **	.35 **
Empathic Concern Δ	.07	.00	-.09	-.10	.11	.24 ⁺
Mirroring T1	.15	.18	-.04	-.04	.00	.06
Mirroring T2	.35 **	.30 *	-.14	.31 *	.33 **	.05
Mirroring Δ	.13	.06	-.07	.31 *	.30 *	.01
<i>MET</i>						
Cognitive Empathy T1	.18	.20	-.04	-.13	-.04	.12
Cognitive Empathy T2	.02	.01	-.03	-.24 ⁺	-.11	.18
Cognitive Empathy Δ	-.17	-.19	.01	-.20	-.11	.13
Emotional Empathy T1	.29 *	.12	-.26 *	.08	.18	.15
Emotional Empathy T2	.25*	.13	-.19	.17	.22 ⁺	.09
Emotional Empathy Δ	.01	.06	.05	.11	.02	-.11
<i>Public Goods Game</i>						
Given Amount in € T1	-.03	.14	.19	-.20	-.17	.02
Given Amount in € T2	-.06	.14	.23 ⁺	-.05	.03	.10
Given Amount in € Δ	-.01	.10	.13	.07	.14	.10
<i>DCS</i>						
Empathic Statements T1	.11	.22 ⁺	.09	-.11	-.09	.04
Empathic Statements T2	.40 **	.23 ⁺	-.26 *	.00	.07	.08
Empathic Statements Δ	.16	-.03	-.23 ⁺	.07	.11	.04
Empathic Feedback T1	.09	.15	.04	-.09	-.19	-.15
Empathic Feedback T2	.17	.11	-.09	-.03	-.05	-.04
Empathic Feedback Δ	.06	-.07	-.14	.04	.12	.10

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

T1=First Session, T2=Fourth Session, PT=Perspective Taking, MP=Mental State Perception.

Table A.21

Partial Correlations Between the CEEQ-GROUP Subscales Empathic Concern and Mirroring and the Remaining Empathy Measures

Measures	EC T1	EC T2	EC Δ	MR T1	MR T2	MR Δ
<i>MET</i>						
Cognitive Empathy T1	.36 **	.42 ***	.11	.17	.12	-.06
Cognitive Empathy T2	.22 ⁺	.31 *	.13	.03	-.07	-.08
Cognitive Empathy Δ	-.08	.00	.09	-.13	-.23 ⁺	-.05
Emotional Empathy T1	.39 **	.37 **	.00	.20	.38 **	.11
Emotional Empathy T2	.43 ***	.43 ***	.03	.16	.37 **	.13
Emotional Empathy Δ	.18	.16	.00	.06	.10	.01
<i>Public Goods Game</i>						
Given Amount in € T1	-.05	.01	.09	.03	-.13	-.17
Given Amount in € T2	.03	.02	.01	.01	-.07	-.11
Given Amount in € Δ	.10	.02	-.08	.05	.02	-.05
<i>DCS</i>						
Empathic Statements T1	.10	.10	.02	.29 *	.06	-.26 *
Empathic Statements T2	.15	.28 *	.17	.03	.05	.02
Empathic Statements Δ	.01	.10	.10	-.22 ⁺	-.03	.22 ⁺
Empathic Feedback T1	.03	.02	.02	-.02	-.21 ⁺	-.17
Empathic Feedback T2	-.12	-.05	.06	.06	-.06	-.11
Empathic Feedback Δ	-.15	-.10	.01	.08	.12	.02

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

T1=First Session, T2=Fourth Session, EC=Empathic Concern, MR=Mirroring.

Table A.22

Partial Correlations Between the MET Subscales Cognitive and Emotional Empathy and the Remaining Empathy Measures

Measures	COG T1	COG T2	COG Δ	EMO T1	EMO T2	EMO Δ
<i>MET</i>						
Cognitive Empathy T1						
Cognitive Empathy T2	.74 ***					
Cognitive Empathy Δ	.00	.67 ***				
Emotional Empathy T1	.20	.11	-.07			
Emotional Empathy T2	.32 **	.24 ⁺	-.02	.75 ***		
Emotional Empathy Δ	.24 ⁺	.20	.01	-.10	.57 ***	

(table continues)

Table A.22 (continued)

Measures	COG T1	COG T2	COG Δ	EMO T1	EMO T2	EMO Δ
<i>Public Goods Game</i>						
Given Amount in € T1	.26 *	.36 **	.25 *	-.09	-.03	.12
Given Amount in € T2	.08	-.13	-.30 *	-.16	-.11	.05
Given Amount in € Δ	-.08	-.35 **	-.45***	-.07	-.09	-.06
<i>DCS</i>						
Empathic Statements T1	.15	.21 ⁺	.16	.05	-.03	-.05
Empathic Statements T2	.47 ***	.33 **	.01	-.04	.01	.05
Empathic Statements Δ	.20	.06	-.13	-.07	.02	.08
Empathic Feedback T1	.17	.11	-.01	.05	.07	.05
Empathic Feedback T2	-.25 *	-.18	.03	.19	-.01	-.25 *
Empathic Feedback Δ	-.45***	-.29 *	.08	.17	-.06	-.33 **

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

T1=First Session, T2=Fourth Session, COG=Cognitive Empathy, EMO=Emotional Empathy.

Table A.23

Partial Correlations Between the Public Goods Game and the DCS Discussion Parameters

Measures	PGG T1	PGG T2	PGG Δ	ES T1	ES T2	ES Δ	EF T1	EF T2
<i>Public Goods Game</i>								
Amount in € T1								
Amount in € T2	.35**							
Amount in € Δ	-.36**	.72***						
<i>DCS</i>								
Emp. Statements T1	.26*	.06	-.11					
Emp. Statements T2	.04	-.15	-.15	.07				
Emp. Statements Δ	-.18	-.14	.00	-.75***	.60***			
Emp. Feedback T1	.14	-.03	-.11	.14	.03	-.09		
Emp. Feedback T2	-.09	-.24 ⁺	-.14	.13	.00	-.09	.56***	
Emp. Feedback Δ	-.27*	-.26 *	-.05	.02	-.02	-.03	-.41**	.52***

Note. ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). $N=91$.

T1=First Session, T2=Fourth Session, PGG=Public Goods Game, ES=Empathic Statements, EF=Empathic Feedback.

8.4 Experimenter Manual

Original German Wording (Blue=Verbal Instructions; Red=Tasks)

ABLAUF DER ERSTEN SITZUNG

[Alle Testleiter begrüßen die Studienteilnehmer zusammen in Raum JK33/213, in der Zeit, in der die Versuchspersonen die Fragebögen am Computer ausfüllen, wertet ein Testleiter die Ergebnisse des Spiels aus]

Begrüßung

[Namen der Studienteilnehmer auf der Teilnehmerliste abhaken]

[jedem Studienteilnehmer sein ID-Kärtchen geben]

[Beginn der Testsitzung auf Sitzungsprotokoll eintragen]

Willkommen zur ersten Sitzung im Rahmen der Studie „Interaktion in Gruppen“. Mein Name ist [...] und das sind meine Kollegen [...] und [...], wir werden die heutige Testsitzung mit Ihnen durchführen. Falls Sie ein Handy dabei haben, würden wir Sie bitten es auszuschalten.

Ablauf der Studie im Allgemeinen

Bevor wir beginnen, möchte ich Ihnen kurz etwas über den allgemeinen Ablauf der Studie „Interaktion in Gruppen“ erzählen. Die Studie umfasst insgesamt vier Sitzungen, die alle in dieser jetzigen Gruppenzusammenstellung stattfinden werden – es sei denn, jemand von Ihnen fällt aus. Am Ende der vierten Sitzung wird Ihre Teilnahme mit 60,- Euro in bar vergütet.

Ihre Daten werden vertraulich behandelt, getrennt von Ihrem Namen aufbewahrt, und ausschließlich zu wissenschaftlichen Zwecken verwendet. Sie können die Sitzung jederzeit abbrechen, daraus würden Ihnen keinerlei Nachteile entstehen und Sie würden anteilig für die heutige Sitzung mit 5 Euro entschädigt werden.

Ich würde Sie jetzt bitten, die Einverständniserklärung vor Ihnen sorgfältig durchzulesen und zu unterschreiben.

Falls Sie Fragen haben, können Sie die jederzeit stellen.

[Einverständniserklärungen einsammeln]

Ablauf der heutigen ersten Sitzung

In der heutigen Sitzung wird es zuerst darum gehen, dass Sie ein paar Fragebögen ausfüllen. Dann sollen sich alle Studienteilnehmer gegenseitig kennenlernen. Danach wird ein gemeinsames Spiel gespielt, bei dem Sie zusätzlich zur Aufwandsentschädigung etwas Geld gewinnen können. Danach folgt eine 10-minütige Gruppendiskussion, die auf Video aufgenommen werden wird. Zum Schluss werden wir Sie bitten, einige weitere Fragebögen am Computer zu beantworten, dazu werden wir den Raum wechseln.

Alles wird noch ausführlich erklärt werden, wenn es soweit ist.

Heute in einer Woche findet dann die zweite Sitzung statt. Darin werden einige Übungen durchgeführt, Spiele gespielt und gemeinsam mit den anderen Studienteilnehmern Aufgaben gelöst, ähnlich wie in der dritten und vierten Sitzung.

Demographischer Fragebogen und IRI

Als nächstes möchten wir Sie bitten, drei kurze Fragebögen auszufüllen, die auf Ihren Plätzen schon verteilt sind. Bitte tragen Sie Ihre ID Nummer auf jedem Fragebogen ein.

[Demographische Fragebögen und IRI einsammeln]

Kennenlernspiel

Vielen Dank!

Als erste Aufgabe der heutigen Sitzung möchten wir Sie bitten, sich jeweils einen anderen Studienteilnehmer zu suchen und 2 Minuten miteinander zu sprechen. Das Thema dürfen Sie sich aussuchen. Die Zeit wird gestoppt. Am Ende der 2 Minuten unterschreiben Sie sich bitte gegenseitig Ihre Zettel, die wir jetzt austeilen.

[Vitamin-C-Zettel austeilen]

Danach suchen Sie sich bitte jemand anderen und sprechen wiederum 2 Minuten miteinander. Das wird so lange fortgesetzt bis jeder mit jedem gesprochen hat.

Die Zettel behalten Sie!

Viel Spaß, es geht jetzt los.

[Die Zeit stoppen, nach 2 Minuten zum Wechsel auffordern bis alle durch sind. Bei ungerader Gruppengröße muss ein Studienteilnehmer immer Pause machen.]

Einführung in die Zielbedingung (nur Experimentalgruppe 2!)

In den nächsten Wochen werden Sie an einem gemeinsamen Gruppenziel arbeiten: Ihre Gruppenaufgabe besteht darin, in den folgenden Sitzungen eine kreative Collage aus Zeitschriften und anderen Materialien zum Thema „Emotionen“ zu erstellen. Das Ziel für Sie als Gruppe besteht darin, uns das Ergebnis ihrer gemeinsamen Arbeit in der letzten Sitzung zu präsentieren. Wir werden Ihnen zu Ihrer gemeinsamen Collage dann eine Rückmeldung geben. Es besteht auch die Möglichkeit, die Collage nach Ende der Studie in den Aufenthaltsräumen einer Klinik aufzuhängen. Diese meist kahlen und weißen Wände würden damit verschönert und Ihre Collage könnte zur emotionalen Anregung der Patienten dienen. Konkret wird Ihre gemeinsame Aufgabe sein, mit der Collage 5 bis 10 verschiedene Emotionen darzustellen. Beispielsweise könnten Sie die unterschiedlichen Komponenten von Emotionen abbilden, d.h. eine Emotion darstellen mithilfe der Gefühle, Gedanken, körperlichen Reaktionen und Verhaltensweisen, die damit einhergehen können. Außerdem wäre es schön, wenn die Collage möglichst bunt und visuell ansprechend ist. Lassen Sie Ihrer Kreativität freien Lauf und probieren Sie die verschiedenen Materialien aus.

Egal, wie sie die Collage gemeinsam gestalten, seien Sie sich bitte stets bewusst, dass es in den nächsten Wochen auf Ihr Teamwork ankommt und wie gut Sie als Gruppe Ihr gemeinsames Ziel erreichen.

[Präsentation der Collage-Materialien (max. 5 min)]

Damit Sie einen Eindruck davon erhalten, wie und womit Sie die Collage in den nächsten Wochen gestalten können, zeigen wir Ihnen bereits heute die Materialien, die Ihnen dabei zur Verfügung stehen werden. Sie können sich den Moderationskoffer und die Zeitschriften einmal kurz zusammen anschauen und vielleicht auch schon gemeinsam erste Ideen sammeln.

[Verteilung von Take-Home-Karten]

Ihr gemeinsames Ziel und eine kurze Beschreibung der konkreten Aufgabe haben wir für Sie auch noch einmal auf diesen Karten hier festgehalten, die Sie gerne mitnehmen können. Nächste Woche können Sie dann mit der Collage beginnen.

[Nachbereitung: Materialien wieder zusammenräumen]

Public-Goods Game

Als nächstes werden wir ein Spiel spielen.

Dazu werden wir Ihnen eine ausführliche Instruktion austeilen, in der das Spiel erklärt wird. Lesen Sie sich diese Instruktion bitte genau durch.

Danach folgen drei Beispiele für einen Durchgang. Falls Sie danach keine Fragen zum Spielablauf haben, geht es los! Um die Anonymität Ihrer Entscheidung zu gewährleisten, möchten wir Sie bitten Ihre 4 Euro mitzunehmen, kurz den Raum zu verlassen und immer nur einzeln diesen Raum wieder zu betreten. Hier haben Sie dann ungestört die Möglichkeit, einen von Ihnen gewählten Teilbetrag in die Gemeinschaftskasse zu legen. Den Rest behalten Sie!

Ihr Gewinn wird Ihnen am Ende der vierten Sitzung ausgezahlt und ist unabhängig von Ihren 60,- Euro Aufwandsentschädigung.

[Instruktion und für jeden 20 x 20 Cent austeilen und Zeit zum Lesen geben, Kasse bereithalten]

[drei Beispiele erläutern: 1. Man behält seine 4 Euro während alle anderen ihre 4 Euro in die Kasse geben; 2. Man gibt seine 4 Euro in die Kasse während alle anderen ihre 4 Euro behalten; 3. Jeder gibt 2 Euro ab; je nach Vpn-Anzahl unterschiedlich]

[Wenn keine Fragen bestehen, Zeit zum Nachdenken geben und Umschläge mit der Kasse einsammeln]

Gruppendiskussion

Vielen Dank!

Als nächstes wird eine Gruppendiskussion zu einem vorgegebenen Thema stattfinden. Das Thema wird gleich eingeblendet. Obwohl die Diskussion aufgenommen wird und Sie sich untereinander noch nicht so gut kennen, möchten wir Sie bitten, offen zu diskutieren. Die Aufnahme wird nur von den an dieser Studie beteiligten Wissenschaftlern gesichtet werden. Im Laufe der Diskussion dürfen und sollen Sie Ihre persönliche Meinung zu dem Thema äußern und auch die Meinung der anderen Teilnehmer bewerten und kritisieren.

Das Thema der Diskussion ist *[Power-Point]* das Rauchverbot. Damit Sie ungestört sind, werden alle Versuchsleiter den Raum verlassen und nach 10 Minuten wiederkommen.

[Kamera einschalten- Anlage „IKD-Voraussetzungen für Videoaufnahmen“ beachten, alle Versuchsleiter verlassen den Raum]

[Nach 10 Minuten wiederkommen, die Diskussion und die Aufnahme stoppen]

Vielen Dank!

Diese Diskussion wird in der vierten und letzten Sitzung noch einmal für 10 Minuten fortgeführt werden.

Arbeit am Computer

Sie haben nun über die Hälfte der ersten Sitzung geschafft.

Als letzte Aufgabe möchten wir Sie bitten, einige letzte Fragebögen zu beantworten, teilweise am Computer. Dazu müssen wir in einen anderen Raum. Ihre Sachen können Sie mitnehmen, von da aus können Sie direkt nach Hause gehen.

[Studienteilnehmer in Testräume im Keller bringen, jedem Teilnehmer den Platz mit seiner ID zuweisen]

MET, CEEQ Group, Fragebogen zu Gruppenprozessen, Gruppenzusammensetzung und Manipulation Check Ziel

Als letzte Aufgabe für heute möchten wir Sie bitten, einige letzte Fragebogen am Computer auszufüllen. Um ein entsprechendes Feld anzukreuzen, verwenden sie bitte ein kleines „x“. Bei einem dieser Fragebögen werden Ihnen zum Teil Fragen gestellt, die sich auf die Gruppe von Studienteilnehmern beziehen, mit denen Sie heute an der Sitzung teilnehmen. Falls die in den Fragen beschriebenen Situationen noch nicht stattgefunden haben sollten, versuchen Sie sich bitte die entsprechende Situation vorzustellen und anhand dieser Vorstellung hypothetisch zu antworten.

Bitte öffnen Sie nacheinander die Excel-Dateien und minimieren Sie diese, wenn Sie mit der Bearbeitung fertig sind. Wir speichern am Ende alles ab. Wenn Sie keine weiteren Fragen haben, können Sie mit dem ersten Fragebogen - dem MET - beginnen. Vergessen Sie bitte nie, ihre ID-Nummern einzuschreiben.

[Reihenfolge: MET, CEEQ Group, Fragebogen zu Gruppenprozessen, Gruppenzusammensetzung und Manipulation Check Ziel]

Danksagung und Beantwortung von Fragen

[für jeden Teilnehmer einzeln, sobald er mit den Fragebögen fertig ist, in einem gesonderten Raum]

Vielen Dank für Ihre Teilnahme! Wir hoffen es hat Ihnen Spaß gemacht. Haben Sie noch irgendwelche Fragen oder Anmerkungen?

[Fragen beantworten, aber nicht zu viel verraten, notfalls auf Debriefing am Ende der vierten Sitzung verweisen]

Verabschiedung und Terminerinnerung

Ihr nächster Termin ist dann am ... um ... Falls Ihnen etwas dazwischen kommen sollte, geben Sie uns bitte frühzeitig bescheid. Auf Wiedersehen!

[Terminkärtchen überreichen]

[Falls danach gefragt wird.]

Wie bei allen Studien werden bei der Studie „Interaktion in Gruppen“ Namen und Daten getrennt. Allerdings werden ID und Name an wenigen Stellen zusammengebracht, was auf den ersten Blick wie eine Datenschutzverletzung aussieht. Diese Stellen sind in dieser Studie folgende:

- Bei den Sitzungen notieren wir auf Teilnahmelisten, wer erschienen ist und welche ID zugeteilt wurde.

Diese Listen sind die Grundlage für den Eintrag in unsere Studien-Datenbank und helfen uns, bei Folgesitzungen die gleiche ID zu verteilen.

Nach der Sitzung werden die Listen eingeschlossen.

- Wie immer gibt es eine Studien-Datenbank im Computer, in welche die Adressen der Studienteilnehmer, Termin und ID eingetragen werden.

Das ist aufgrund des Datenschutzgesetzes nötig, nach welchem wir garantieren müssen, dass “jeder ST zu jeder Zeit die Vernichtung seiner Daten fordern” kann. In so einem Fall müssten wir zurückverfolgen können, welcher Datensatz gelöscht werden soll. Die Datenbank ist nur für projektinterne Mitarbeiter zugänglich und durch ein Passwort geschützt.

ID-Nr. Systematik (NICHT den Versuchspersonen mitteilen)

Die IDs werden systematisch vergeben. Sie beinhalten Informationen über die Untersuchungsbedingung (1. Ziffer) und das Geschlecht des Studienteilnehmers (2. Ziffer). Die letzten drei Ziffern sind fortlaufend von 001 bis 999 für jeden Studienteilnehmer durchnummeriert.

Die erste Ziffer kann also nur Werte zwischen 1 und 3 annehmen, die zweite nur Werte zwischen 1 und 2.

1. Versuchsbedingung	1 = Kontrollgruppe 2 = Experimentalgruppe 1 3 = Experimentalgruppe 2
2. Geschlecht	1 = männlich 2 = weiblich
3. / 4. / 5.	fortlaufend 001-999

ABLAUF DER ZWEITEN SITZUNG

[Alle Testleiter begrüßen die Studienteilnehmer zusammen in Raum JK33/213, die komplette Testsitzung wird dort stattfinden]

Begrüßung und Ablauf der heutigen Sitzung

[Namen der Studienteilnehmer auf der Teilnehmerliste abhaken]

[Beginn der Testsitzung auf Sitzungsprotokoll eintragen]

Willkommen zur zweiten Sitzung im Rahmen der Studie „Interaktion in Gruppen“. Mein Name ist [...] und das sind meine Kollegen [...] und [...], wir werden die heutige Testsitzung mit Ihnen durchführen. Falls Sie ein Handy dabei haben, würden wir Sie bitten es auszuschalten.

Die heutige Sitzung wird aus drei Teilen bestehen, von denen jeder ca. eine halbe Stunde dauern wird. Im ersten werden Sie zusammen Filmausschnitte sehen und diese im Hinblick auf darin gezeigte Emotionen und Kognitionen diskutieren. Im zweiten werden Sie ein Spiel spielen, bei dem es darum geht dargestellte Emotionen zu erraten. Am Schluss werden Sie an einem Projekt arbeiten, das noch näher erläutert wird und bei dem Sie kreativ werden können. In den letzten fünf Minuten werden wir Sie bitten, ein bis zwei kurze Fragebögen auszufüllen.

Falls Sie Fragen haben, können Sie die jederzeit stellen.

MASC-Übung

Als erste Aufgabe der heutigen Sitzung möchten wir Ihnen einige Filmausschnitte zeigen, die Sie diskutieren sollen. Lesen Sie sich dazu die eingeblendeten Instruktionen bitte aufmerksam durch.

[Power-Point Präsentation ggf. vorlesen und erläutern]

Falls Sie keine weiteren Fragen haben, geht es jetzt los.

[Moderation: Bei jedem MASC-Item die Gruppe diskutieren lassen, notfalls nach 1. Minute eine Einigung auf eine Antwortalternative erzwingen. Ggf. zwischendurch daran erinnern, dass auch persönliche Erfahrungen als Argumentationsquelle herangezogen werden können und sollen. Rückmelden, ob Antwort richtig oder falsch war.]

[Falls nach 30 Minuten der MASC noch nicht abgeschlossen ist, abbrechen und mit dem Spiel fortfahren]

Vielen Dank!

Spiel: Emotionen erraten

Als nächstes möchten wir ein Spiel mit Ihnen spielen. Es ist ähnlich wie Tabu, falls das jemand von Ihnen kennt. Dazu möchten wir Sie in zwei Gruppen einteilen.

[nach der Sitzordnung so einteilen, dass keine zwei Spieler direkt nebeneinandersitzen, die zusammenspielen; bei ungerader Anzahl von Teilnehmern z.B. eine 2er und eine 3er Gruppe einteilen]

Das Spiel beginnt, nachdem ein Spieler aus einer der zwei Gruppen verdeckt eine dieser Karten hier zieht. Auf diesen Karten stehen Emotionen. Nun ist es an dem Spieler, diese Emotion darzustellen. Er darf dabei nicht sprechen oder aufstehen, darf aber Gesten zur Unterstützung verwenden oder seine nächste Umgebung mit einbeziehen. Wenn es den Mitspielern seiner eigenen Gruppe gelingt, die Emotion auf der Karte richtig zu erraten, erhält die Gruppe einen Punkt und er zieht eine neue Karte – falls das Erraten nicht klappt, darf der Spieler auch aufgeben und eine neue Karte ziehen. Das geht so lange weiter, bis 1 Minute um ist; danach ist ein Spieler aus der anderen Gruppe dran *[bei gleich großen Gruppen im Uhrzeigersinn]*.

Haben Sie dazu noch Fragen?

[Spiele durchführen, Zeit stoppen, Punkte notieren; bis 30 Minuten um sind.]

Vielen Dank!

Projektarbeit

KG UND EG1

Ihre nächste Aufgabe ist es, in Einzelarbeit eine kreative Collage aus Zeitschriften und anderen Materialien zum Thema „Emotionen“ zu erstellen. Konkret besteht Ihre Aufgabe darin, 3 bis 5 verschiedene Emotionen darzustellen. Welche Emotionen Sie wählen, können Sie sich gerne selbst frei aussuchen. Bitte nutzen Sie alle Ihnen zur Verfügung stehenden Materialien – Bilder und Texte in Zeitschriften, bunte Karten und Stifte – und lassen Sie Ihrer Kreativität und Ihren Ideen freien Lauf. Beispielsweise könnten Sie die unterschiedlichen Komponenten von Emotionen abbilden, d.h. eine Emotion darstellen mithilfe der Gefühle, Gedanken, körperlichen Reaktionen und Verhaltensweisen, die damit einhergehen können. Außerdem wäre es schön, wenn die Collage möglichst bunt und visuell ansprechend ist.

Sie haben heute 25 min. Zeit für die Collage und auch nächste Woche noch die Möglichkeit weiter daran zu arbeiten. In der letzten Sitzung werden wir Sie dann darum bitten, sich Ihre Collagen gegenseitig kurz vorzustellen. Es besteht auch die Möglichkeit, die Collage nach Ende der Studie in den Aufenthaltsräumen einer Klinik aufzuhängen. Diese meist kahlen und weißen Wände würden damit verschönert und Ihre Collage könnte zur emotionalen Anregung der Patienten dienen.

[Nach 30 min. beenden]

Vielen Dank, nächste Woche können Sie mit der Arbeit an Ihren Collagen fortfahren.

EG2

Wie in der letzten Woche angekündigt und auf Ihren Karten festgehalten, werden Sie heute mit der Arbeit an Ihrem gemeinsamen Gruppenziel beginnen. Noch einmal zu Ihrer Erinnerung: Ihre Gruppenaufgabe besteht darin, in den nächsten Sitzungen eine kreative Collage aus Zeitschriften und anderen Materialien zum Thema „Emotionen“ zu erstellen. Das Ziel für Sie als Gruppe besteht darin, uns das Ergebnis ihrer gemeinsamen Arbeit in der letzten Sitzung zu präsentieren. Es besteht außerdem die Möglichkeit, die Collage nach Ende der Studie in den Aufenthaltsräumen einer Klinik aufzuhängen, wo sie dann zur emotionalen Anregung von Patienten dienen könnte. Konkret besteht Ihre Aufgabe darin, 5 bis 10 verschiedene Emotionen mithilfe der Collage darzustellen. Welche Emotionen Sie zusammen als Gruppe wählen und präsentieren, können Sie sich gerne selbst frei aussuchen. Bitte nutzen Sie alle Ihnen zur Verfügung stehenden Materialien – Bilder und Texte in Zeitschriften, bunte Karten und Stifte – und lassen Sie Ihrer Kreativität und Ihren Ideen freien Lauf. Beispielsweise könnten Sie die unterschiedlichen Komponenten von Emotionen

abbilden, d.h. eine Emotion darstellen mithilfe der Gefühle, Gedanken, körperlichen Reaktionen und Verhaltensweisen, die damit einhergehen können. Außerdem wäre es schön, wenn die Collage möglichst bunt und visuell ansprechend ist.

Sie haben 25 min. Zeit für die Gruppenaufgabe. Egal, wie sie die Collage gemeinsam gestalten, seien Sie sich bitte stets bewusst, dass es auf Ihr Teamwork ankommt und wie gut Sie als Gruppe Ihr gemeinsames Ziel erreichen.

[Nach 30 min. beenden]

Vielen Dank, nächste Woche können Sie mit der Arbeit an Ihrer gemeinsamen Collage fortfahren.

Fragebogen zu Gruppenprozessen und zur Gruppenaufgabe (letzterer nur EG 2!)

Als Letztes möchten wir Sie noch bitten den ausgeteilten Fragebogen zu beantworten (EG 2: die zwei ausgeteilten Fragebögen zu beantworten). Bitte tragen Sie Ihre ID ein.

Danksagung und Beantwortung von Fragen

Vielen Dank für Ihre Teilnahme! Wir hoffen es hat Ihnen Spaß gemacht. Haben Sie noch irgendwelche Fragen oder Anmerkungen?

[Fragen beantworten, aber nicht zu viel verraten, notfalls auf Debriefing am Ende der vierten Sitzung verweisen]

Verabschiedung und Terminerinnerung

Ihr nächster Termin ist dann am ... um ... Falls Ihnen etwas dazwischen kommen sollte, geben Sie uns bitte frühzeitig bescheid. Auf Wiedersehen!

[Terminkärtchen überreichen]

ABLAUF DER DRITTEN SITZUNG

[Alle Testleiter begrüßen die Studienteilnehmer zusammen in Raum JK33/213, die komplette Testsitzung wird dort stattfinden]

Begrüßung und Ablauf der heutigen Sitzung

[Namen der Studienteilnehmer auf der Teilnehmerliste abhaken]

[Beginn der Testsitzung auf Sitzungsprotokoll eintragen]

Willkommen zur dritten Sitzung im Rahmen der Studie „Interaktion in Gruppen“. Mein Name ist [...] und das sind meine Kollegen [...] und [...], wir werden die heutige Testsitzung mit Ihnen durchführen. Falls Sie ein Handy dabei haben, würden wir Sie bitten es auszuschalten. Die heutige Sitzung wird aus vier Teilen bestehen. Im ersten wird Ihnen ein Ausschnitt eines Filmes präsentiert. Im zweiten werden Sie, jeder von Ihnen für sich, eine kurze Erinnerungs-Übung durchführen. Danach folgen Gespräche in Zweiergruppen. Und schließlich werden wir Sie bitten die Collage aus der letzten Sitzung fertig zu stellen.

Falls Sie Fragen haben, können Sie die jederzeit stellen.

Filmvorführung

Am Anfang werden wir Ihnen einen Filmausschnitt zeigen. Es handelt sich dabei um die letzten 15 Minuten des Films „das Streben nach Glück“ mit Will Smith. Falls Sie den Film noch nicht kennen, möchten wir Sie bitten sich die Handlung des Films, die wir für Sie auf einer Seite zusammengefasst haben, durchzulesen.

Mit der Filmvorführung ist eine kleine Aufgabe verbunden. Wir möchten Sie bitten, auf die Mimik, Gestik und das Verhalten des Protagonisten zu achten (Will Smith alias Chris Gardner). Schreiben Sie dabei bitte alle Emotionen auf dem vor Ihnen liegenden Schreibblock auf, die Sie zu erkennen glauben. Das ist nicht immer einfach.

Falls es Ihnen schwer fällt, die Emotion einzugrenzen oder Sie mehrere auf einmal zu erkennen glauben, versuchen Sie sich trotzdem für eine zu entscheiden. Sie können dazu z.B. die Emotionswörter benutzen, die im Spiel von letzter Woche vorgekommen sind.

Falls Sie keine weiteren Fragen haben, geht es jetzt los.

*[Film abspielen]*Erinnerungs-Übung

Als Nächstes möchten wir Sie bitten, sich an fünf positive Lebensereignisse in Ihrem Leben zu erinnern. Das müssen nicht so entscheidende Ereignisse sein, wie eben im Film – aber sie sollten für Sie emotional bedeutsam sein. Weiterhin sollten die Ereignisse möglichst nicht intim sein. Beispiele wären z.B. das gute Abschneiden in einer wichtigen Prüfung, eine schöne Reise, ein beruflicher Erfolg, o.ä.

Bitte schreiben Sie diese Ereignisse auf.

Weiterhin möchten wir Sie bitten, sich so genau wie möglich an die Emotionen zu erinnern, die Sie bei jedem dieser Ereignisse empfunden haben, und diese ebenfalls zusammen mit dem Ereignis aufzuschreiben. Auch hier können Sie dazu die Wörter aus dem Spiel von letzter Woche benutzen (z.B. stolz, erleichtert, glücklich, fröhlich).

Sie haben für das Erinnern und Aufschreiben insgesamt 10 Minuten Zeit.

Falls Sie keine weiteren Fragen haben, geht es jetzt los.

Gespräche in Zweiergruppen

Im dritten Teil dieser Sitzung möchten wir Sie bitten, sich jeweils einen anderen Studienteilnehmer zu suchen. In den Zweiergruppen sollten sich beide Partner von je einem ihrer positiven Ereignisse und den dabei empfundenen Emotionen berichten. Nach ca. 5 Minuten wird gewechselt, und über ein anderes Ereignis sowie die entsprechenden Emotionen geredet. Das Ganze verläuft wie beim Kennenlernspiel in der ersten Sitzung so lange, bis jeder mit jedem gesprochen hat.

Die Versuchsleiter werden zu den Gesprächen den Raum verlassen und alle fünf Minuten kurz hereinkommen, um das Signal zum Wechseln zu Geben.

Falls die Gruppe nicht vollständig aufgeteilt werden kann, muss einer von Ihnen pausieren; nach Möglichkeit nicht zweimal dieselbe Person hintereinander.

Falls Sie keine weiteren Fragen haben, geht es jetzt los.

[Die Versuchsleiter verlassen den Raum, alle 5 Minuten zum Wechseln auffordern]

Projektarbeit

KG UND EG1

Wie auch letzte Woche, werden Sie auch heute an Ihrer Collage weiterarbeiten und sie fertig stellen. Sie haben auch dieses Mal wieder 25 min. Zeit dafür.

[Nach 30 Minuten beenden]

Vielen Dank, nächste Woche können Sie sich Ihre fertigen Collagen dann gegenseitig präsentieren.

EG2

Wie auch letzte Woche, werden Sie auch heute an Ihrem gemeinsamen Ziel weiterarbeiten und Ihre Collage fertig stellen. Sie haben auch dieses Mal wieder 25 min. Zeit dafür.

Bitte seien Sie sich dabei bewusst, dass es auch heute wieder auf Ihr Teamwork ankommt und wie gut Sie als Gruppe Ihr gemeinsames Ziel erreichen.

[Nach 30 Minuten beenden]

Vielen Dank, nächste Woche können Sie das Ergebnis Ihrer gemeinsamen Arbeit präsentieren. Wir werden Ihnen dann eine kurze Rückmeldung dazu geben.

Fragebogen zu Gruppenprozessen und zur Gruppenaufgabe (letzterer nur EG 2!)

Als Letztes möchten wir Sie noch bitten den ausgeteilten Fragebogen zu beantworten (EG 2: die zwei ausgeteilten Fragebögen zu beantworten). Bitte tragen Sie Ihre ID ein.

Danksagung und Beantwortung von Fragen

Vielen Dank für Ihre Teilnahme! Wir hoffen es hat Ihnen Spaß gemacht. Haben Sie noch irgendwelche Fragen oder Anmerkungen?

[Fragen beantworten, aber nicht zu viel verraten, notfalls auf Debriefing am Ende der vierten Sitzung verweisen]

Verabschiedung und Terminerinnerung

Ihr nächster Termin ist dann am ... um ... Falls Ihnen etwas dazwischen kommen sollte, geben Sie uns bitte frühzeitig bescheid. Auf Wiedersehen!

[Terminkärtchen überreichen]

ABLAUF DER VIERTEN SITZUNG

[Alle Testleiter begrüßen die Studienteilnehmer zusammen in Raum JK33/213, wenn das Public-Goods Game beginnt geht ein Testleiter in den Keller und richtet in der Zeit des Spiels die Computer ein. Nach dem Spiel bleibt ein Testleiter und wertet die Ergebnisse des Spiels aus]

Begrüßung

[Namen der Studienteilnehmer auf der Teilnehmerliste abhaken]

[Beginn der Testsitzung auf Sitzungsprotokoll eintragen]

Willkommen zur letzten vierten Sitzung im Rahmen der Studie „Interaktion in Gruppen“. Mein Name ist [...] und das sind meine Kollegen [...] und [...], wir werden die heutige Testsitzung mit Ihnen durchführen. Falls Sie ein Handy dabei haben, würden wir Sie bitten es auszuschalten.

Ablauf der heutigen Sitzung

Der Verlauf der heutigen Sitzung wird ziemlich genau der ersten Sitzung dieser Studie entsprechen. Zuerst werden wir Sie bitten Ihre Collage(n) vorzustellen. Danach werden wir dasselbe Spiel spielen, bei dem Sie wieder etwas Geld gewinnen können. Danach wird nochmals eine zehnminütige Gruppendiskussion stattfinden, die aufgenommen werden wird, und am Schluss werden wir Sie bitten, einige Fragebögen am Computer im Keller zu beantworten.

Vorstellung der Collagen

Wir möchten Sie jetzt bitten sich gegenseitig Ihre Collagen vorzustellen. Bitte erläutern Sie dabei kurz, welche Emotionen Sie dargestellt haben.

Experimentalgruppe 2: Wir haben nun noch einen kurzen Fragebogen für Sie.

[Fragebogen zur Gruppenarbeit / Collage austeilen]

Bitte vergessen Sie nicht, Ihre ID einzutragen.

Public-Goods Game

Als Nächstes werden wir dasselbe Spiel spielen, das Sie ganz am Anfang der Studie gespielt haben. Zur Erinnerung haben wir Ihnen die Instruktion noch einmal ausgeteilt. Bitte lesen Sie sich diese noch einmal durch.

Wir werden Ihnen nochmals kurz Beispiele geben. Falls Sie danach keine Fragen zum Spielablauf haben, geht es los! Ihr Gewinn wird Ihnen heute am Ende der Sitzung ausgezahlt und ist unabhängig von Ihren 60,- Euro Aufwandsentschädigung.

[Instruktion und für jeden 20 x 20 Cent austeilen und Zeit zum Lesen geben, Kasse bereithalten]

[drei Beispiele erläutern: 1. Man behält seine 4 Euro während alle anderen ihre 4 Euro in die Kasse geben; 2. Man gibt seine 4 Euro in die Kasse während alle anderen ihre 4 Euro behalten; 3. Jeder gibt 2 Euro ab]

[Wenn keine Fragen bestehen, Zeit zum Nachdenken geben und Umschläge mit der Kasse einsammeln]

Gruppendiskussion

Vielen Dank!

Als Nächstes wird die zweite Gruppendiskussion stattfinden, die auch diesmal aufgenommen wird. Hierzu auch wieder der Hinweis, dass Sie im Laufe der Diskussion Ihre persönliche Meinung zu dem Thema äußern sollen und auch die Meinung der anderen Teilnehmer bewerten und kritisieren dürfen.

Das Thema ist dabei dasselbe wie in der ersten Sitzung, wobei heute auch etwas allgemeiner über das Rauchen an sich diskutiert werden kann. Als Anstoß für die Diskussion können Sie folgende Zitate benutzen.

[Zitate einblenden]

Damit Sie ungestört sind, werden alle Versuchsleiter den Raum verlassen und nach 10 Minuten wiederkommen.

Wir möchten Sie bitten anzufangen! Die Aufnahme beginnt jetzt.

[Kamera einschalten- Anlage „IKD-Voraussetzungen für Videoaufnahmen“ beachten, alle Versuchsleiter verlassen den Raum]

[Nach 10 Minuten wiederkommen, die Diskussion und die Aufnahme stoppen]

Vielen Dank!

Arbeit am Computer

Sie haben nun über die Hälfte der letzten Sitzung geschafft.

Als letzte Aufgabe möchten wir Sie bitten, einige Fragebögen am Computer zu beantworten. Dazu müssen wir wieder in die Computerräume im Keller. Ihre Sachen können Sie mitnehmen, von da aus können Sie direkt nach Hause gehen.

[Studienteilnehmer in Testräume im Keller bringen, jedem Teilnehmer einen Platz zuweisen]

Fragebögen am Computer

MET, CEEQ Group, Fragebogen zu Gruppenprozessen, Gruppenzusammensetzung, Manipulation Check Ziel, SES, IRI, NEO-FFI

Als letzte Aufgabe für heute möchten wir Sie bitten, einige letzte Fragebogen am Computer auszufüllen. Um ein entsprechendes Feld anzukreuzen, verwenden sie bitte ein kleines „x“. Bei einem dieser Fragebögen werden Ihnen zum Teil Fragen gestellt, die sich auf die Gruppe von Studienteilnehmern beziehen, mit denen Sie heute an der Sitzung teilnehmen. Falls die in den Fragen beschriebenen Situationen noch nicht stattgefunden haben sollten, versuchen Sie sich bitte die entsprechende Situation vorzustellen und anhand dieser Vorstellung hypothetisch zu antworten.

Bitte öffnen Sie nacheinander die Excel-Dateien und minimieren Sie diese, wenn Sie mit der Bearbeitung fertig sind. Wir speichern am Ende alles ab. Wenn Sie keine weiteren Fragen haben, können Sie mit dem ersten Fragebogen - dem MET - beginnen. Vergessen Sie bitte nie, ihre ID-Nummern einzuschreiben.

[Reihenfolge: MET, CEEQ Group, Fragebogen zu Gruppenprozessen, Gruppenzusammensetzung, Manipulation Check Ziel, SES, IRI, NEO-FFI]

Debriefing und Beantwortung von Fragen

[für jeden Teilnehmer einzeln, sobald er mit dem NEO-FFI fertig ist, in einem gesonderten Raum]

Vielen Dank für Ihre Teilnahme! Wir hoffen es hat Ihnen Spaß gemacht.

[Debriefing „Worum es in dieser Studie ging“ aushändigen und alle aufkommenden Fragen beantworten, bei Interesse an Studienergebnissen Ende 2011 in Mailingliste aufnehmen]

[Studienteilnehmer darum bitten, die Inhalte des Debriefings mit Blick auf die Hauptstudie vorerst für sich zu behalten]

Auszahlung der Spielgewinne und der Aufwandsentschädigung

[Restbetrag des Gewinns des Public-Goods Games aus der ersten und letzten Sitzung auszahlen und gesamten Gewinn QUITTIEREN lassen]

[60,- Euro auszahlen und QUITTIEREN lassen]

8.5 Debriefing

Original German Wording

Worum es in dieser Studie ging

Der Hauptgegenstand der Studie, an der Sie teilgenommen haben, ist die Erforschung von Empathie in Kleingruppen. Konkret ging es um die Frage, ob bestimmte *Gruppenparameter* einen Einfluss auf Empathie haben und in welche Richtung diese wirken können.

Es gab dabei drei unterschiedliche Arten von Gruppen, denen die Studienteilnehmer zugewiesen wurden. Die erste Gruppe war eine Kontrollgruppe, die zweite Gruppe war so wenig divers wie möglich (alle Teilnehmer hatten dasselbe Geschlecht und Alter und dieselbe Muttersprache), die dritte Gruppe hatte ein gemeinsames Ziel, was die Gruppenmitglieder verbinden sollte. Im Laufe der vier Sitzungen sollten dabei Unterschiede zwischen den drei Gruppen hinsichtlich der Veränderung in ihrer Empathie untersucht werden.

Die erste und die vierte Sitzung diente dabei gänzlich der Messung von Empathie mit psychologischen Methoden, sowohl ganz allgemein als auch konkret gegenüber den Gruppenmitgliedern. Als direkte und indirekte Maße für Empathie verwendeten wir dabei Fragebögen, bildbasierte Tests, Diskussionsparameter und die Tendenz, innerhalb der Geldabgabe-Spiele im Sinne der Gruppe zu handeln. Die zweite und dritte Sitzung beinhaltete eine Art Kurzintervention zum Steigern von Empathie. Wir vermuten, dass diese Intervention sich je nach Gruppe unterschiedlich ausgewirkt hat, z.B. dass sich die Zunahme von Empathie in der zweiten und dritten Gruppe als größer erweisen wird.

Was ist der Sinn einer solchen Forschungsarbeit? Ohne die Fähigkeit zur Empathie, dem Fühlen für Andere, wäre menschliches Zusammenleben sicherlich schwer vorstellbar - und nicht nur weniger leistungsfähig sondern sicherlich auch weniger wertvoll. In vielen Kontexten wäre die Steigerung von Empathie hilfreich und wünschenswert. Unsere Forschung bemüht sich darum, geeignete Methoden dafür zu finden. Die Dynamik innerhalb einer Gruppe positiv auszunutzen, könnte möglicherweise eine davon sein.

Wir hoffen, Sie hatten Spaß beim Experiment und bedanken uns herzlich für Ihre Teilnahme!

English Translation

What this study was about

The main objective of the study you have been participated in is to investigate empathy in small groups. More concretely, it focused the question if specific group parameters have an impact on empathy and if they are positively or negatively related to empathy.

The participants of this study were divided into three different categories of groups. The first group was a control group, the second had as low levels of diversity as possible (all participants had the same gender, age, and native language), and the third group had a shared goal, which was supposed to bind the group members to each other. Group differences in the development of empathy were meant to be investigated throughout the four sessions.

The first and fourth session entailed a measurement of empathy with psychological methods, including general levels of empathy and specific levels of empathy towards the group members. Questionnaires, image-based tests, discussion parameters and prosocial orientation within the economic game have been used as direct and indirect measures of empathy. The second and third session entailed a kind of short-term intervention with the objective to enhance empathy. We suppose that this intervention differently affected the three categories of groups, e.g. that the level of empathy will be increased within the second and the third group.

What is the purpose of such an investigation? It is difficult to imagine human coexistence without empathy, the ability to feel with other persons – human coexistence would be not only less efficient but also less valuable without it. Within many situations, an enhancement of empathy would be helpful and desirable. Our research aims to find appropriate methods to realize this objective. It is possible that the positive usage of group dynamics represents one of these methods.

We hope that you enjoyed this experiment and thank you very much for your participation!

List of Figures

Figure 1.	Interdependence of Cognitive Functions and Processes within Cognitive Empathy	9
Figure 2.	Interdependence of Emotional Functions and Processes within Emotional Empathy	10
Figure 3.	Interdependence of Cognitive and Emotional Empathy	11
Figure 4.	Within-Partner and Between-Partner Correlation Matrix, Based on Gonzales & Griffin (1996, 1997)	41
Figure 5.	The Actor-Partner Interdependence Model (APIM).....	43
Figure 6.	The Experimental Procedure of the Present Study	67
Figure 7.	Factor Analysis of Empathy Measures in the First Experimental Session	114
Figure 8.	Factor Analysis of Cognitive Empathy Measures in the First Experimental Session.....	115
Figure 9.	Factor Analysis of Emotional Empathy Measures in the First Experimental Session.....	115
Figure 10.	Factor Analysis of Empathy Measures in the Fourth Experimental Session.....	117
Figure 11.	Factor Analysis of Cognitive Empathy Measures in the Fourth Experimental Session	118
Figure 12.	Factor Analysis of Emotional Empathy Measures in the Fourth Experimental Session	118
Figure 13.	Marginal Means for the IRI Subscale Perspective Taking.....	120
Figure 14.	Marginal Means for the IRI Subscale Fantasy	120
Figure 15.	Marginal Means for the IRI Subscale Empathic Concern	121
Figure 16.	Marginal Means for the IRI Subscale Personal Distress	121
Figure 17.	Marginal Means for the CEEQ-GROUP Subscale Perspective Taking	123
Figure 18.	Marginal Means for the CEEQ-GROUP Subscale Mental State Perception	123
Figure 19.	Marginal Means for the CEEQ-GROUP Subscale Empathic Concern	124
Figure 20.	Marginal Means for the CEEQ-GROUP Subscale Mirroring.....	124

Figure 21.	Marginal Means for the MET Subscale Cognitive Empathy	126
Figure 22.	Marginal Means for the MET Subscale Emotional Empathy	126
Figure 23.	Marginal Means for the Results of the Public Goods Game	127
Figure 24.	Marginal Means for the DCS Parameter Empathic Content.....	129
Figure 25.	Marginal Means for the DCS Parameter Empathic Feedback	129
Figure 26.	Mediator Effect of Static Group Climate on the Interrelationship Between the Shared Goal Condition and the Effects of the Public Goods Game Within the Fourth Session	132
Figure 27.	Hypothesized Overlap Between Empathy, Openness, and General Intelligence	162

List of Tables

Table 1.	Constructs Related to Empathy: Definitions and Differentiating Features	17
Table 2.	Composition of the Experimental Groups at the Beginning of the Investigation Period	57
Table 3.	Demographic Characteristics of the Total Sample and the Three Experimental Conditions	59
Table 4.	Descriptive Statistics of Empathy Measures for the First Experimental Session.....	87
Table 5.	Difference Between Means and Standard Deviations in the First and Fourth Session	88
Table 6.	Reliability of Empathy Measures.....	90
Table 7.	Means and Standard Deviations of the NEO-FFI	92
Table 8.	Means and Standard Deviations of the Questionnaire on Group Processes	93
Table 9.	Separate Manipulation Check for all Facets of Perceived Surface-Level Diversity.....	97
Table 10.	Means and Standard Deviations of the Questionnaire on Task Motivation and Orientation	102
Table 11.	Reliability of the Questionnaire on Task Motivation and Orientation	103
Table 12.	Moderated Regression Testing for a Moderating Effect of Changes in Group Climate on the Interrelationship Between Experimental Condition and Changes in the Results of the Public Goods Game.....	130
Table 13.	Moderated Regression Testing for a Moderating Effect of Static Group Climate on the Interrelationship Between Experimental Condition and Static Results of the Public Goods Game in the Fourth Experimental Session.....	131
Table 14.	Moderated Regression Testing for a Moderating Effect of Changes in Group Identification on the Interrelationship Between Experimental Condition and Changes in the Results of the Public Goods Game	133
Table 15.	Moderated Regression Testing for a Moderating Effect of Changes in Perceived Diversity on the Interrelationship Between Experimental Condition and Changes in the Results of the Public Goods Game	134
Table 16.	Moderated Regression Testing for a Moderating Effect of Static Perceived Diversity on the Interrelationship Between Experimental Condition and Changes in the Results of the Public Goods Game.....	135

Table 17.	Moderated Regression Testing for a Moderating Effect of Changes in Perceived Diversity on the Interrelationship Between Experimental Condition and Static Results of the Public Goods Game	136
Table 18.	Moderated Regression Testing for a Moderating Effect of Static Perceived Diversity on the Interrelationship Between Experimental Condition and Static Results of the Public Goods Game	136

List of Tables in Appendix

Table A1.	Operationalization of Empathic Statements.....	205
Table A2.	Operationalization of Empathic Feedback.....	205
Table A3.	Bivariate Correlations Between Social Desirability, Big Five Personality Traits, and Empathy in the First Experimental Session.....	206
Table A4.	Bivariate Correlations Between Social Desirability, Big Five Personality Traits, and Empathy in the Fourth Experimental Session	207
Table A5.	Bivariate Correlations Between Social Desirability, Big Five Personality Traits, and ΔM in Empathy Measures Between the First and the Fourth Experimental Session	208
Table A6.	Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation, and Empathy (First Session)	209
Table A7.	Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation, and Empathy (Fourth Session)	210
Table A8.	Bivariate Correlations Between ΔM in Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation, and ΔM in Empathy.....	211
Table A9.	Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation in First Session and Empathy in Fourth Session	212
Table A10.	Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation in Fourth Session and Empathy in First Session	213
Table A11.	Bivariate Correlations Between ΔM in Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation, and Empathy in the First Session.....	214
Table A12.	Bivariate Correlations Between ΔM in Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation, and Empathy in the Fourth Session.....	215
Table A13.	Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation in First Session and ΔM in Empathy	216

Table A14.	Bivariate Correlations Between Group Climate, Group Identification, Dominance, Affiliation, Perceived Diversity, Diversity Beliefs, Goal Representation in Fourth Session and ΔM in Empathy	217
Table A15.	Associations Between Age, Gender, Education, and Empathy in the First Session	218
Table A16.	Associations Between Age, Gender, Education, and Empathy in the Fourth Session.....	219
Table A17.	Associations Between Age, Gender, Education, and ΔM in Empathy Measures	220
Table A18.	Partial Correlations Between the IRI Subscales Perspective Taking and Fantasy and the Remaining Empathy Measures.....	221
Table A19.	Partial Correlations Between the IRI Subscales Empathic Concern and Personal Distress and the Remaining Empathy Measures.....	222
Table A20.	Partial Correlations Between the CEEQ-GROUP Subscales Perspective Taking and Mental State Perception and the Remaining Empathy Measures	223
Table A21.	Partial Correlations Between the CEEQ-GROUP Subscales Empathic Concern and Mirroring and the Remaining Empathy Measures.....	224
Table A22.	Partial Correlations Between the MET Subscales Cognitive and Emotional Empathy and the Remaining Empathy Measures.....	224
Table A23.	Partial Correlations Between the Public Goods Game and the DCS Discussion Parameters.....	225