

Best Practices in Reference Checking: How SkillSurvey's Pre-Hire 360® Measures-Up

A best practice in reference checking involves **360-degree (multi-source) feedback** on the candidate's work behavior in his/her previous jobs. It is based upon the premise that the prediction of future work behavior can be **predicted by feedback** from those individuals -- managers, peers, direct reports, and clients -- who have worked with the candidate in the past. This paper discusses *360 degree* feedback, as well as other best practices in reference checking, and provides **statistical evidence** to address how the SkillSurvey Pre-Hire 360® measures up.

Introduction

In order to improve quality-of-hire, and to retain top-performing individuals, an employer needs to go beyond the materials provided by the job candidate, such as the resume, cover letter and employment application. Personality test results can provide additional information about the candidate, such as innate tendencies towards assertiveness, persuasion or conscientiousness. Yet, these test results are also gathered through candidate self-reports. And, although this information provided by the candidate gives the employer a sense of the candidate's training and experience, or even of the candidate's motivational qualities, none of these materials provides any indication as to the candidate's competence, or ability, with respect to the behaviors required for success on the job.

But, this is exactly what every employer wants to know. They want the answer to the question, "Does this candidate have the competencies needed not only to succeed, but also to excel, in the position that I am trying to fill within my organization?" This is in essence getting at quality-of-hire.

If done well during the hiring process, reference checking can help to answer the question of whether a candidate has the competencies required for job success, and thus whether he or she would be considered a 'quality hire.'

Best Practices for a Reference-Checking Process That Improves Quality-of-Hire and How SkillSurvey Measures Up

Below are the criteria for best practices for a reference-checking process. It is important that organizations evaluate their current process against these criteria.

1. **Assess competency in behaviors that are valid or critical to work success.**

Competencies are considered to be enduring characteristics of an individual. This definition is in alignment with the repeated finding that past behavior is a good predictor of future performance.

The word 'competence' is typically associated with the degree to which someone accomplishes an activity, or performs at work. If a candidate displays competence in a work activity, he/she is considered to exhibit strength in that area. Lack of competence in an area may not be a deal-breaker, but it certainly indicates an area where the candidate has room for developmental opportunity. A competency assessment, when administered before bringing a candidate on-board, can help to identify which candidate has the greatest potential for success on the job. Furthermore, this practice can help to identify areas to target when designing the new employee's on-boarding program.

SkillSurvey follows best practices in creating valid competency clusters and individual competency items. These best practices include gathering information about the job and the organizational culture by conducting focus

groups, administering surveys, or holding one-on-one conversations with key stakeholders. In some instances, subject matter experts are called upon to provide additional expertise. In other cases, SkillSurvey collaborates with the customer and incorporates validated competencies from the company's previous efforts in the areas of on-boarding and performance management. Using a seven-point point rating scale, and a 360-degree format, all references are asked to rate the candidate's competencies on the validated work behaviors. Text boxes are provided for entry of verbatim comments on the candidate's strengths and areas for developmental improvement. Candor is achieved, as all feedback is aggregated into a final report; thus, specific feedback from any one reference is not associated with that individual.

SkillSurvey works with clients to conduct validity (outcome) studies. In fact, our recent validity studies have revealed that reference feedback gathered with SkillSurvey's Pre-Hire 360® was statistically predictive of retention. Additionally, a study that also included supervisor evaluations of those hired showed SkillSurvey's Pre-Hire 360® to be statistically predictive of those ratings.

2. Strive for consistency and reliability.

Strive for a consistent, reliable process. If the organization chooses not to follow the same reference-checking procedure for candidates at all levels, then try to put guidelines in place to follow the same procedure for all candidates *for a given position*. This process is not only fair to both employer and candidate, but it also helps to mitigate against the risk of litigation.

Reliability also has a specific meaning with respect to tests and assessments. In general, reliability of any assessment is the degree to which the results obtained are consistent. There are two primary measures of reliability: internal consistency and test-retest reliability.

Results are highly consistent -- across all surveys, Cronbach's alpha ranges from .96 to .98.

Internal consistency is the extent to which all items in a scale measure the same construct,

such as competence at work. The statistical index of internal consistency is called *Cronbach's alpha*. SkillSurvey has created over 200 validated surveys for reference checking. The library consists of surveys created for a particular job family or industry, as well as surveys customized for a particular job at a given company. Across the many surveys, Cronbach's alpha ranges from .96 - .98. These coefficients indicate that the surveys in the SkillSurvey library demonstrate very high internal consistency in their measurement of competence at work, in a variety of jobs.

Test-retest reliability is gathered when an assessment is administered to a sample of individuals on two separate occasions, and results are produced on both of these occasions. A correlation coefficient is usually computed to assess the relationship, or consistency, between results gathered at Time 1, compared to results gathered at Time 2. The SkillSurvey Pre-Hire 360® demonstrates excellent test-retest reliability. After 2-3 weeks had elapsed since the first rating, all references in a given time period were asked to provide a second set of ratings on a candidate. Over 40% of all references complied with this request, yielding 2,974 sets of ratings at both Time 1 and Time 2. The consistency, or reliability between the two sets of ratings, was highly statistically significant, achieving a correlation coefficient of +.764 ($p < .0001$). Both managers and non-managers exhibited a statistically high degree of consistency in their ratings ($r = +.79$ and $r = +.72$, respectively).

References are extremely consistent in the feedback that they provide on job candidates. In a sample of 2,974 candidates, the correlation coefficient between feedback at Time 1 vs. Time 2 was highly statistically significant ($r = +.764$; $p < .0001$).

3. Make sure that your process is free from discrimination and adverse impact.

Research studies have shown that reference checking, like other materials gathered on job candidates, can result in adverse impact. One recent finding was of gender bias in the contents of letters of recommendation for candidates applying for professional positions. Furthermore, other studies focusing on job incumbents (employees) have reported both gender and race/ethnicity bias in supervisor ratings of performance.

SkillSurvey monitors the potential for adverse impact with respect to gender, age, and race/ethnicity in the Pre-Hire 360® feedback collection and reporting. The optional form for the collection of race/ethnicity data is in alignment with the recent EEO-1 form (U.S. Department of Labor, Office of Federal Contract Compliance Programs, or OFCCP).

► Bottom Line for SkillSurvey Reference Checking ◀

Using the 4/5 rule, there was no evidence of adverse impact with respect to race/ethnicity, gender, or age group, meaning that SkillSurvey's Pre-Hire 360® supports our clients' efforts to be ADEA, EEOC and OFCCP compliant.

Methods. Using the principles outlined in the Department of Labor's *Uniform Guidelines on Employee Selection Procedures*, SkillSurvey collects demographic information on job candidates. On an optional research form, job candidates are asked to respond to four questions, one each for race/ethnicity group, gender, birth cohort, and age group. Recent analyses of a random sample of 150,000 candidates revealed that 85.2% had answered these demographic items. Data from this sample were used for adverse impact analyses.

Results. Presence of adverse impact was assessed by applying the 4/5 rule. The *Uniform*

Guidelines states that, "A selection rate for any race, sex, or ethnic group which is less than 4/5 (80%) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact."

Note that SkillSurvey does not compute or report a selection rate, *per se*. SkillSurvey does, however, report an overall score for each candidate. The majority of candidates in the sample (84.9%) received an overall score in the "High/Very High" category. It is membership in this category that was used for all tests to assess presence of adverse impact.

Table 1. No race/ethnicity adverse impact with SkillSurvey reference checking, in compliance with the U.S. Equal Employment Opportunity Commission (EEOC).

	n	% High/ Very High	Meets 4/5 Rule
White	91,234	85.1	n/a
Af Am	15,523	84.9	YES
Asian	8,514	83.5	YES
Hisp	8,816	85.7	YES

Race/ethnicity grouping. A group of 91,234 candidates indicated that they were white. Of this group, 85.1% received a "High/Very High" overall score when reference checked using SkillSurvey. In the sample were also 15,523 Black/African Americans, 8,514 Asians, and 8,816 Hispanics. The percentages in these race/ethnic groups who received "High/Very High" scores were as follows: African American (84.9%); Asian (83.5%); and Hispanic (85.7%). As depicted in Table 1, using the 4/5 rule, there was no adverse impact for race/ethnicity.

Table 2. No gender adverse impact with SkillSurvey reference checking, in compliance with the U.S. Equal Employment Opportunity Commission (EEOC).

	n	% High/ Very High	Meets 4/5 Rule
M	45,587	83.7	n/a
F	86,273	85.7	YES

Gender. A group of 45,587 candidates indicated that they were male (M). Of this group, 83.7% received a "High/Very High" overall score when reference checked with SkillSurvey. A total of 86,273 candidates indicated that they were female (F), and 85.7% of this group received a "High/Very High" overall score when reference checked with SkillSurvey. As depicted in Table 2, using the 4/5 rule, there was no adverse impact for gender.

Table 3. No age group adverse impact with SkillSurvey reference checking, in compliance with the U.S. Age Discrimination in Employment Act (ADEA).

	n	% High/Very High	Meets 4/5 Rule
<=39 yrs	87,951	85.3	n/a
>=40 yrs	40,528	85.1	YES

Age group. A group of 87,951 candidates indicated that they were less than 40 years old. Of this group, 85.3% received a "High/Very High" overall score when reference checked with SkillSurvey. A total of 40,528 candidates indicated that they were 40 years old or older, and these candidates received a "High/Very High" overall score 85.1% of the time. The data in Table 3 support the claim that there was no adverse impact for age group, based on application of the 4/5 rule.

Position breakouts. Statistical analyses were performed on random samples of candidates applying for different positions (e.g., administrative, hourly, general professional, I.T., nursing, and executive). The results indicated that there was no adverse impact for gender, race/ethnicity, or age group within the different groups of candidates.

4. Gather multiple perspectives on the job candidate – from managers, peers, direct-reports, and clients who have worked with the candidate.

The majority of respondents (85%) in a 2004 SHRM study reported that someone from HR is most likely to provide a reference on a job candidate who they were considering. It could be argued, however, that in most cases the candidate did not work with anyone in HR. At the very least, someone who has managed the

candidate should provide feedback on the candidate's work performance. After all, the manager is most knowledgeable as to the performance goals and expectations set for the candidate, and whether these goals were met. Yet, in order to reach individual, team, and organizational goals, the candidate must competently work with individuals in other roles at the company. For example, the candidate must often work closely and effectively with peers and other coworkers in order to successfully complete cross-functional initiatives. The candidate may manage others, and therefore play a critical role in recruiting, coaching, and retaining top-performing employees. For these reasons, and others, gathering what is called 'multi-source' or '360-degree' feedback on a candidate's past work performance, is essential.

Analysis of recent SkillSurvey data – a random sample of 150,000 job candidates and 624,433 references -- revealed that when a 360-degree feedback procedure was used for reference checking, an 84.7% response rate was achieved, on average. This translated to 4.16 *references per job candidate*. Most of the references (53.3%) were from managers (including teachers for entry-level candidates). Because of the "360" format, references were also gathered from business peers (38.6%), direct-reports (3.4%) or clients (4.7%).

5. Ensure that you get candid feedback from your reference providers by providing them anonymity.

From the employer's perspective, it is critical that accurate and candid feedback be gathered on the job candidate, a feature that is missing from most of today's reference-checking processes. But, what about the benefits of candid feedback to the job candidate, once he or she is hired? Candid feedback has the potential to improve both organizational and individual development. Numerous research studies have linked accurate self-perception with successful work outcomes. Yet, individuals differ greatly in their ability to see themselves as others see them. Candid feedback, when shared with an employee, and incorporated into an individual developmental plan, can reap great benefits for individuals, teams, and organizations. Because some may feel uncomfortable with the thought of providing candid feedback, especially to others in the workplace, it is very important to

implement steps into your process that provide anonymity to those providing the feedback.

On average, SkillSurvey's reference-checking process yields an 84.7% reference response rate, which translates to 4.16 references per candidate.

As mentioned earlier, the average reference response rate of the SkillSurvey Pre-Hire 360® is 84.7%. In addition, a recent analysis of verbatim comments from 2,000 references showed that the majority of all references provided substantive verbatim comments as to the candidate's specific strengths (78.1% of all references) and areas in which they could improve (69.7% of all references).

The majority of SkillSurvey references provide substantive, verbatim comments regarding the candidate's specific work-related strengths and areas for improvement.

6. Embrace automation.

The cost and time savings that come with automating the reference-checking process are enormous. Some methods of reference checking (letters of recommendation or telephone reference checks) can take up to two weeks to complete. Automation of reference checking with an on-demand process that leverages the Internet and mobile technology has reduced the turn-around time to within two business days, thereby reducing the time spent reference checking by 92%. And because references often complete the reference check on days when they are out of the office (e.g., travel days, weekends, etc.), or after traditional work hours, the reference check may be available even earlier. This reduces the time to fill a position and thus translates into more resources for other strategic HR initiatives.

Using the SkillSurvey on-demand process, references respond, on average, within two business days, reducing the time spent reference checking (by phone) by 92%.

Other benefits of automation include the elimination or reduction of data-entry errors. If done well, automation can assist with providing a consistent and documented process for all candidates. Finally, automation permits more sophisticated methods of analysis, such as an earlier evaluation of trends in your recruiting and hiring processes.

7. Document your process and results.

It is vitally important that you document and appropriately store the materials that you collect from job candidates. In a worst-case scenario, such as an audit or legal claim, thorough documentation and accessibility of these records will prove invaluable. It is often the case that results from telephone reference checks are recorded in writing, and then stored in multiple paper files. In contrast, all reference-checking feedback gathered by SkillSurvey is automatically stored on a secure central server.

This document provides insight regarding best practices for reference checking your job candidates. For additional information regarding SkillSurvey's Pre-Hire 360®, an online pre-hire competency assessment, visit www.blr.com/skillsurvey.

Resources

- Anastasi, A. (1988). *Psychological testing*. New York: Macmillan Publishing Company.
- Bliss, W. (2001). *Legal, effective references: How to give and get them*. Alexandria, VA: SHRM.
- Brafman, O., & R. Brafman. (2008). *Sway: The irresistible pull of irrational behavior*. NY: Doubleday.
- Cronbach, L.J. (1960). *Essentials of psychological testing (2nd ed.)*. Oxford, England: Harper.
- Dipboye, R.L. (1994). Structured and unstructured selection interviews: Beyond the job-fit model. In G. R. Ferris (Ed.), *Research in personnel and human resources management* (Vol. 12, pp. 79-123). Greenwich, CT: JAI Press.
- Fernandez-Araoz, C. (2007). *Great people decisions: Why they matter so much, why they are so hard, and how you can master them*. NJ: John Wiley & Sons.
- Goldsmith, M. (2007). *What got you here won't get you there*. NY: Hyperion.
- Goleman, D. (1998). *Working with emotional intelligence*. NY: Bantam Books.
- Guarino, A.C. (2007). *Smart is not enough!* NY: John Wiley & Sons.
- Guion, R.M., (1998). *Assessment, measurement, and prediction for personnel decisions*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hedricks, C., Kessler, S., & Rupayana, D. (April, 2011). *Candidate motivation and multi-source feedback can identify quality nurses*. Accepted for presentation at the 26th annual meeting of the Society for Industrial and Organizational Psychology, Chicago, IL.
- Hedricks, C.A., Oswald, F., & Robie, C. (April, 2009) *Reliability and validity of on-demand 360° reference checking*. Presented at the 24th annual meeting of the Society for Industrial and Organizational Psychology, New Orleans, LA.
- Hedricks, C.A. & Robie, C. (April, 2008). *Online reference checking reduces group differences*. . Presented at the 23rd annual meeting of the Society for Industrial and Organizational Psychology, San Francisco, CA.
- Hough, L.M., Oswald, F.O., & Ployhart, R.E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment*, 9, 152-194.
- Landy, F. J. (Ed.). (2005). *Employment discrimination litigation: Behavioral, quantitative, and legal perspectives*. San Francisco: Jossey-Bass.
- Lencioni, P. (2002). *The five dysfunctions of a team*. San Francisco, CA: Jossey-Bass.
- Levashina, J., & Campion, M.A. (2007). Measuring faking in the employment interview: Development and validation of an interview faking behavior scale. *Journal of Applied Psychology*, 92, 1638-1656.
- Madera, J., Hebl, M., & Martin, R. (2009). Gender and letters of recommendation for academics: Agentic and communal differences. *Journal of Applied Psychology*, 94, 1591-1599.
- Netemeyer, R.G., Maxham III, J.G. (2007). Employee versus supervisor ratings of performance in the retail customer service sector: Differences in predictive validity for customer outcomes. *Journal of Retailing*, 83, 131-145.
- Nunnally, J.C. and Bernstein, I.H. (Eds.) (1994). *Psychometric theory (3rd ed)*. New York: McGraw-Hill.
- O'Connor, R.M., & Little, I.S. (2003). Revisiting the predictive validity of emotional intelligence: Self report versus ability-based measures. *Personality and Individual Differences*, 35, 1893-1902.
- O*Net (2000). *Testing and assessment: An employer's guide to good practices*. U.S. Department of Labor Employment and Training Administration [online]. Available: www.onetcenter.org/dl_files/empTestAsse.pdf

- Outtz, J.L. (2010). *Adverse impact: Implications for organizational staffing and high stakes selection*. NY: Routledge.
- Pfeffer, J., & Sutton, R.I. (2006). *Hard facts, dangerous half-truths, and total nonsense: Profiting from evidence-based management*. Boston, MA: Harvard Business School Press.
- Pfeffer, J., & Sutton, R.I. (2000). *The knowing-doing gap*. Boston, MA: Harvard Business School Press.
- Ryan, A.M., & Lasek, M. (1991). Negligent hiring and defamation: Areas of liability related to pre-employment inquiries. *Personnel Psychology*, 44, 293-319.
- Schmidt, F.L., & Hunter, J.E. (1998). The validity and utility of selection models in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.
- Schoorman, F.D., & Mayer, R.C. (2008). The value of common perspectives in self-reported appraisals: You get what you ask for. *Organizational Research Methods*, 11, 148-159.
- SkillSurvey, Inc. (2009). *Health system increases quality of hire. Improves retention. Studies validate 360° reference checking*. Wayne, PA: SkillSurvey, Inc. [collaborative study with Poudre Valley Health System, Colorado]
- Society for Human Resource Management. (2005). *2004 Reference and background checking survey report*. Alexandria, VA: SHRM.
- Society for Industrial and Organizational Psychology (SIOP). (2003). *Principles for the validation and use of personnel selection procedures (4th ed.)*. Society for Industrial and Organizational Psychology [online]. Available: www.siop.org/_Principles/principles.pdf
- Spencer, L.M., & Spencer, S.M. (1993). *Competence at work: Models for superior performance*. NY: John Wiley & Sons.
- Stolz, R.F. (May 16, 2007). Taking the lead (cover story). *Human Resources Executive* magazine.
- U.S. Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, and Department of Justice. (1978). Uniform guidelines on employee selection procedures. *Federal Register*, 43, 38290-38315.
- Welch, J. & Welch, S. (2005). *Winning*. NY: HarperCollins.
- Williamson, L.G., Campion, J.E., Malos, S.B., Roehling, M.V., & Campion, M.A. (1997). Employment interview on trial linking interview structure with litigation outcomes. *Journal of Applied Psychology*, 82, 900-912.
- Wood, R., & Payne, T. (1998). *Competency based recruitment and selection: A practical guide*. NY: John Wiley & Sons.
- Zimmerman, R.D., Triana, M.C., & Barrick, M.R. (2010). Predictive criterion-related validity of observer ratings of personality and job-related competencies using multiple raters and multiple performance criteria. *Human Performance*, 23, 361-378
-
- About the author:** Cynthia A. Hedricks, Ph.D. (Chief Analytics Officer, SkillSurvey, Inc.) has over 25 years of diverse experience collecting, analyzing, and presenting data gathered in the applied setting. In her role at SkillSurvey, Dr. Hedricks integrates market demands with technological HR innovations and psychometric research principles. Before joining SkillSurvey in June of 2007, Cynthia was Senior Vice President of Research & Development for nine years at Caliper, in Princeton, New Jersey.
- Her most recent projects have included: numerous investigations of the reliability, validity and compliance of on-demand reference checking; relationships between employee competence and patient satisfaction in a healthcare setting; a global study of the personality qualities that distinguish women leaders; and the development of norms for cross-cultural comparisons and global talent management. Dr. Hedricks has both national and international consulting experience, helping to improve quality-of-hire at companies such as CH2M HILL, eBay, Citi, Emory Healthcare; Florida Hospitals, GlaxoSmithKline, Help Desk Institute [United Kingdom], Johnson Diversey, Kohl's, Quest Diagnostics, Poudre Valley Health System, Penn Medicine, Strayer University, and

Wal-Mart [Canada, China, and Mexico]. She is a member of the American Society for Healthcare Human Resources Association (ASHHRA), the National Association for Health Care Recruitment (NAHCR), the Society for Human Resources (SHRM) and the Society for Industrial and Organizational Psychology (SIOP, Division 14 of the American Psychological Association). She has conducted research for and has been published in the areas of candidate selection, nursing, occupational science/therapy, and hormones and behavior. She has taught numerous graduate and undergraduate courses, such as research methods, statistics, physiological psychology, and the meaning of work, and most recently at the University of Southern California. She has a Ph.D. in Biopsychology from The University of Chicago.