

Initial Evaluations in the Interview: Relationships with Subsequent Interviewer Evaluations and Employment Offers

Murray R. Barrick and Brian W. Swider
Texas A&M University

Greg L. Stewart
University of Iowa

The authors of this study examine how evaluations made during an early stage of the structured interview (rapport building) influence end of interview scores, subsequent follow-up employment interviews, and actual internship job offers. Candidates making better initial impressions received more internship offers ($r = .22$) and higher interviewer ratings ($r = .42$). As predicted, initial evaluations of candidate competence extend beyond liking and similarity to influence subsequent interview outcomes from the same interviewer ($\Delta R^2 = .05$), from a separate interviewer ($\Delta R^2 = .05$), and from another interviewer who skipped rapport building ($\Delta R^2 = .05$). In contrast, assessments of candidate liking and similarity were not significantly related to other judgments when ratings were provided by different interviewers. The findings of this study thus indicate that initial impressions of candidates influence employment outcomes, and that they may be based on useful judgments of candidate competence that occur in the opening minutes of the structured interview.

Keywords: employment interview, self-presentation tactics, impression management, rapport building, first impression

Face-to-face interviews are frequently used to gather information about job applicants (Jelf, 1999). Research over the past decade has clearly established that the use of structure enhances the job relevance of this information (Campion, Palmer, & Campion, 1997; Huffcutt & Arthur, 1994). Yet structured interviews also contain an unstandardized rapport-building stage where interviewers put candidates at ease by engaging in “small-talk” conversations (Chapman & Zweig, 2005). An unanswered question concerns whether initial impressions from rapport building influence—even interfere with—the collection of useful information during the structured interview.

Research in social psychology suggests that during social greetings, individuals develop perceptions of strangers almost instantaneously and effortlessly (Ambady & Rosenthal, 1992). Indeed, initial impressions may be instinctual, as human beings naturally use whatever information is available to facilitate interaction and survival (Bar, Neta, & Linz, 2006). In the interview context, recent research shows that information communicated through something as simple as a handshake, a smile, or the manner of dress can influence impressions (Barrick, Shaffer, & DeGrassi, 2009; Stewart, Dustin, Barrick, & Darnold, 2008). Thus, one of our purposes in this study was to examine whether “fast and frugal” judgments formed during rapport building, even before any structured questions are asked, affect interview outcomes.

The social interaction inherent in the rapport-building stage can be viewed as a potential threat to collection of job-related information, particularly since the candidate’s primary agenda is to get a job offer. Yet, cognitive researchers have shown that individuals can render judgments of competence quickly. Specifically, Willis and Todorov (2006) found stranger ratings of competence made after a 100-ms exposure to be highly correlated with ratings provided by people familiar with the participant. These findings suggest interviewers may form evaluations of competence even before structured questions are asked, and the impressions formed during rapport building may contain job-relevant information. Yet, research also shows judgments about liking and similarity form quickly (Zajonc, 1980). In fact, Howard and Ferris (1996) suggested that three evaluations are routinely made by interviewers: competence of, affect toward, and similarity to candidates. Since evaluations of liking and similarity in the interview are viewed as irrelevant to job performance (Posthuma, Morgeson, & Campion, 2002), the “biasing” effect of initial impressions on interview outcomes is an important issue. In this study, we assessed the extent to which impressions formed during rapport building capture information beyond liking and similarity.

Research in social psychology also shows surprising levels of consensual agreement in strangers’ social perceptions, even after very brief encounters (Watson, 1989). Employment interviews are usually conducted by a single interviewer, which makes it difficult to separate perceptions made early during rapport building from subsequent evaluations of answers to structured questions. This likely inflates these relationships. Yet, evidence of agreement in stranger ratings implies perceptions of candidate competence formed early in one interview may generalize beyond that candidate–interviewer dyad. Although the magnitude of the correlation should be smaller, finding a meaningful relationship between initial impressions of competence, similarity to, or liking for a candidate from one interviewer and hiring recommendations or

This article was published Online First August 16, 2010.

Murray R. Barrick and Brian W. Swider, Department of Management, Mays Business School, Texas A&M University; Greg L. Stewart, Department of Management, Tippie School of Business, University of Iowa.

Correspondence concerning this article should be addressed to Murray R. Barrick, Department of Management, 420 Wehner, 4221 TAMU, Mays Business School, Texas A&M University, College Station, TX 77843. E-mail: mbarrick@mays.tamu.edu

interview scores provided by others can better illustrate the ubiquitous effect of initial impressions in the interview context.

Given that initial impressions form so quickly, it is also important to explore whether simply not allowing the candidate and interviewer to interact or engage in rapport building before structured questions are asked can reduce or eliminate the formation of initial impressions based on non-job-relevant information. If initial evaluations made during rapport building are unrelated to interview scores on a separate interview without rapport building, then it indicates that limiting interviewer use of rapport building may serve as a way to restrict the impact ancillary information has on the structured interview.

Research is thus needed to allow investigators to (a) understand the effect of initial impressions on employment offers, (b) assess whether initial impressions extend beyond affect and similarity, and (c) explore whether initial impressions generalize beyond a specific interview. In the following three sections, we develop hypotheses to answer questions concerning each of these aspects of the initial impression.

Do General, Overall Initial Evaluations Link to Employment Offers?

Research suggests that decision making occurs in two separate stages (Kahneman, 2003). The first stage is an intuitive judgment resulting in quick categorizations of individuals. The second stage involves more rational processing in which information is carefully weighed and results in revisions to initial categorizations. Recent theoretical treatments assume interviewers also process information in two stages (Dipboye, 2004; Stevens, 2000). The focus of this research has been on more objective pre-interview information such as an application blank or cognitive ability test score. However, another possible source of information that may affect interviewer decision making is judgments made during the social exchange inherent in rapport building.

Unlike information obtained before individuals meet face to face, the quick, intuitive judgments formed during rapport building are not necessarily capturing job-relevant information (Chapman & Zweig, 2005). Although some believe that interviewers make quick hiring decisions on the basis of early subjective evaluations (Anderson, 1960; Springbett, 1958), the empirical research supporting these claims is not well developed (Buckley & Eder, 1988). Yet, on the basis of psychological research highlighting the importance of initial evaluations in social encounters (Ambady & Rosenthal, 1992) and preliminary interview research showing that interviewers make quick decisions (Springbett, 1958), we predicted that the interviewer initial impressions of the candidate correspond with actual employment offers. Therefore,

Hypothesis 1: Initial impression evaluations made before asking structured questions are positively related to subsequent employment offers received by candidates.

Do Initial Evaluations Reflect Perceptions of Competence Beyond Affect and Similarity?

Focusing on the job relevance of early impressions allows us to begin unraveling the mystery about the content upon which initial evaluations are made. While it has been shown that each of the

three primary evaluations (competence, liking, similarity) influence interviewer decision making, they are not equally job-relevant. Specifically, evaluations of liking and similarity are seen as reflecting non-job-relevant information in the interview (Howard & Ferris, 1996; Posthuma et al., 2002). However, research shows people are predisposed to form affective impressions and make judgments of similarity toward a stranger quickly (Ambady & Rosenthal, 1992; Funder & Colvin, 1991). Thus, even if they are not job-related, liking and similarity evaluations should begin to form as soon as the interviewer and candidate meet.

Yet, interviewers are also assumed to make decisions based on perceptions of competence (Howard & Ferris, 1996). Candidates who are evaluated as competent during the initial impression should be viewed as more qualified, receive more favorable interview scores (Howard & Ferris, 1996), and be expected to become higher performers (Huffcutt & Arthur, 1994). But what is the extent to which competence is assessed via initial evaluations? The findings of social cognition researchers confirm the swiftness in which competency evaluations can develop (Willis & Todorov, 2006). Swift evaluations are believed to be based on both physical characteristics (e.g., physical or verbal attractiveness) and perceived personality (Ambady & Rosenthal, 1992), which have each been shown to predict interviewer end-of-interview evaluations of candidate suitability (Barrick et al., 2009; Van Iddekinge, Raymond, & Roth, 2005). Existing findings of a positive relationship between information used to develop initial evaluations (e.g. physical appearance or personality) and ratings of overall interview performance led us to predict that initial evaluations of competence made during rapport building are positively related to structured interview ratings. Indeed, the relationship between assessments of competence and interview outcomes should exist even after variance due to perceptions of liking and similarity has been taken into account. Hence,

Hypothesis 2: Initial evaluations of candidate competence made before structured questions are asked will account for variance in performance during the structured interview beyond the variance accounted for by early evaluations of liking and similarity.

Do the Effects From Initial Evaluations of Candidate Competence Generalize Beyond One Interview?

A full understanding of an initial impression of candidate competence formed during rapport building requires exploration of whether effects extend beyond the specific interviewer to scores from other interviewers, even actual offers for internships and jobs. The next three hypotheses address whether these effects extend beyond one interview, while also ruling out alternative explanations for these findings.

We began this study by asking whether a global, holistic initial impression predicted actual employment offers. This would require the information on which these initial evaluations are based to be relatively stable attributes affecting multiple decision makers (i.e., interviewers). To understand whether these impressions are likely to be job related, we also focused on the content of this impression, specifically whether it is primarily based on competence or instead is largely driven by affect or similarity to the candidate. Prior research (Ambady & Rosenthal, 1992) shows

initial evaluations of competence are based on relatively stable, trait-based information (i.e., personality, physical attractiveness). These attributes have been shown to positively relate to overall ratings of candidate suitability (Barrick et al., 2009; Van Iddekinge et al., 2005). Thus, the influence from perceptions of competence is likely to be reflected in the decisions of others. According to this reasoning, impressions of candidate competence formed initially in one interview are expected to also affect the number of second interview invitations extended from actual employers. In contrast, perceptions of affect and similarity have previously been found to be relatively idiosyncratic to the dyadic relationship between the interviewer and candidate in the one interview (Posthuma et al., 2002). Therefore,

Hypothesis 3: Initial evaluations of candidate competence will account for incremental variance (above and beyond liking and similarity) in the proportion of second interviews extended by actual employers.

To study the effect of initial evaluations, researchers must ask the interviewer about his or her impressions. Yet, the very act of asking for an explicit evaluation may accentuate the influence that the initial impression has on subsequent thoughtful evaluations. To rule out a potential “demand” effect and isolate the intuitive first stage from the rational second stage decision-making process, we obtained a second initial impression rating from a different interviewer who participated in a rapport-building-only interview. Although evidence that the initial impression rating in a mock interview correlates with actual employment decisions helps to minimize concern about a demand effect, we were able to examine specifically whether an initial impression rating in one interview correlated with an evaluation of the same candidate’s responses to structured interview questions in a second interview with a different interviewer. Because prior research has revealed consensus or agreement among individuals when evaluating strangers (Watson, 1989) and because stranger ratings of competence converge with ratings from those highly familiar with the participant (Willis & Todorov, 2006), we hypothesized:

Hypothesis 4: Initial evaluations of candidate competence will account for incremental variance in performance (above and beyond liking and similarity) during a structured interview rated by a different interviewer.

A related question concerns whether the impact of initial impressions of an interviewer based on a social exchange prior to the interviewer’s asking the candidate job-related questions can be reduced simply by forcing the interviewer to begin asking structured questions within seconds of meeting the candidate. Eliminating time for rapport building could possibly inhibit the interviewer from forming impressions prior to eliciting job-relevant information. However, research implying that impressions form almost instantaneously and effortlessly suggests that it may be impossible to keep interviewers from forming early impressions. We explored this possibility by including yet another evaluation from a third interviewer, who was not allowed to interact with the candidate before asking structured questions. The relationship between the structured evaluation provided by an interviewer who simply began asking standardized, job-relevant questions as soon

as the candidate appeared, and an early impression rating formed during rapport building provided by a different rater was then assessed to determine whether the relationship between the early impression and the structured evaluation persisted. On the basis of the notion that impressions form almost instantaneously (Ambady & Rosenthal, 1992), we predicted a significant relationship even though the interviewer did not engage in rapport building. Thus,

Hypothesis 5: Initial evaluations of candidate competence will account for incremental variance in performance (above and beyond liking and similarity) during a structured interview rated by a different interviewer who was not allowed to build rapport with the candidate before asking structured questions.

Method

Participants and Procedure

We collected data from two separate samples over a 2-year period. Participants in both Sample 1 ($N = 76$) and Sample 2 ($N = 113$) were undergraduate students enrolled in a program for professional accountants (PPA). The participants’ mean age was 20.5 years ($SD = 0.69$), and 70% were women. As for the racial and ethnic backgrounds of the participants, 86% were White, 5% were Asian, 4% were Hispanic, and 5% were “other.”

In Sample 1, participants were interviewed in a mock setting that consisted of a 2- to 3-min rapport-building period, followed by a 20- to 25-min structured interview based on 12 job-relevant questions. For these participants, we also tracked subsequent job offers received from actual accounting firms.

Data for Sample 2 were collected 1 year later for a different group of participants. For this sample, each participant went through the same full-length interview used the previous year. Participants in Sample 2 also did two additional interviews. The second was a rapport-building-only interview, which lasted 2–3 min and focused on innocuous, non-job-related questions (e.g., hobbies, weather, and the day’s happenings). The third interview consisted of a structured interview only; that is, the interviewer began asking 12 parallel structured questions as in the full interview within seconds of meeting the candidate. The last two interviews were counterbalanced in order. Sample 2 participants thus took part in three separate mock interviews: a) a full interview, b) a rapport-building-only interview, and c) a structured-portion-only interview. We also assessed interview outcomes for Sample 2 by capturing invitations for a second interview with the firms.

Participants were instructed to treat the mock interview just as they would a “real” interview and were encouraged to dress appropriately, prepare for a formal interview, and so forth. Participants in Sample 2 ($N = 113$) reported a mean score of 4.52 (range from 1 = *strongly disagree* to 5 = *strongly agree*) on three items used to assess whether they approached the mock interview in a serious manner similar to their approach to an actual interview.

Sixty-two human resource majors from a master’s program in the business school conducted the mock interviews. Prior to the interview, each interviewer participated in a 2-hr training session developed around current research on the structured interview (Campion et al., 1997; Posthuma et al., 2002), with a focus on the four dimensions of the structured interview described by Chapman

and Zweig (2005). A summary of the specific job description was also reviewed. Standardized questions and rating scales to be used during all mock interviews were thoroughly discussed. Also, interviewers were instructed to limit any follow-up questions (Campion et al., 1997). When discussing rapport building, which was described as a brief introduction, interviewers were instructed to learn and use the candidate’s name as well as briefly describe their own background. They were instructed not to ask job-relevant questions during this 2- to 3-min introduction. Thus, the mock interview corresponds to an interview the candidates could expect to engage in during their own job search. Interviewers reported a mean of 4.81 (on a scale from 1 = *strongly disagree* to 5 = *strongly agree*) on three items used to assess whether they treated the mock interview the same as an actual interview.

Within 2 weeks after the mock interviews, participants engaged in actual interviews with the “Big Four” accounting firms (Deloitte Touche Tohmatsu, Ernst & Young, KPMG, and PricewaterhouseCoopers). All four firms signed an agreement to interview all PPA students for internship opportunities. This agreement further stipulated when and how long recruiting contact with the candidates could occur. Therefore, the nature and amount of contact between candidates and employers were rigorously standardized. Each Big Four accounting firm extended internship offers to just over half of the PPAs, and all offers were made on one specific date.

Measures

Many of the measures were consistent across the two samples, but some were different. Figure 1 illustrates the data available to test each hypothesis.

Overall initial impressions from rapport building. At the end of the rapport-building stage but prior to starting the structured interview (2–3 min into the interview), interviewers in Sample 1 were asked to complete a perceptual measure of their initial

impression of the candidate. The overall initial impressions evaluation consisted of five questions ($\alpha = .93$) previously used to assess perceived applicant suitability (Cable & Judge, 1997; Higgins & Judge, 2004; Stevens & Kristof, 1995). An example item is “This candidate appears to be very qualified” (response options ranged from 1 = *strongly disagree* to 5 = *strongly agree*).

Initial evaluations after rapport building. Interviewers in both Samples 1 and 2 were asked to complete a perceptual measure of initial evaluations to assess the candidate’s competence, liking for the candidate, and perceived “personal” similarity to the candidate. We used nine items previously developed to assess the three components (Howard & Ferris, 1996). Three items ($\alpha = .88$) were used to assess perceived applicant competence; an example is “This applicant would interact with our clients and other employees quite effectively.” Three items ($\alpha = .85$) were used to assess perceived liking for the candidate; a sample item is “I would like to work with this person.” The third scale—similarity or fit between the interviewer and the candidate—was assessed with three items ($\alpha = .74$) and includes the item “This applicant and I have many of the same beliefs and values.” All questions were rated on a 5-point scale (response options ranged from *strongly disagree* to *strongly agree*).

We conducted a confirmatory factor analysis to assess the discriminant validity of the three scales. Results are reported in Table 1. We tested five competing factor structures: (a) one-factor model in which all nine items were set to load on a global initial impression factor (Model 1), (b) a series of models that constrained scale sets (Models 2–4), and (c) our theoretically driven three-factor model with items for assessment of competence, liking, and similarity loading on separate factors (Model 5). The one-factor model of a global initial impression had poor fit, $\chi^2(27) = 263.02$, $p < .01$; $\chi^2/df = 9.74$; standardized root-mean-square residual (SRMR) = .10; root-mean-square error of approximation

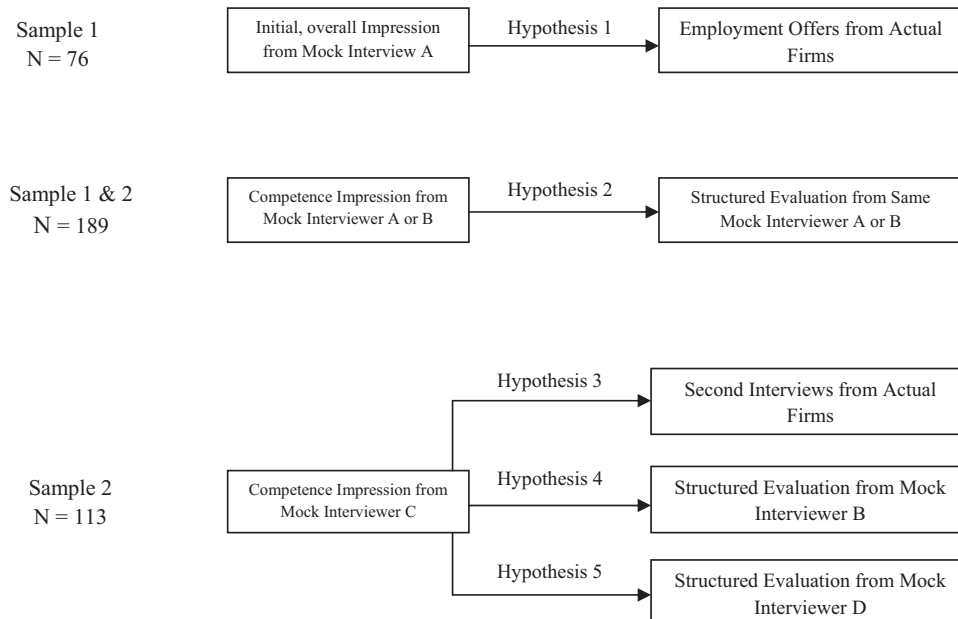


Figure 1. Summary of data collected from each sample of research participants.

Table 1
Confirmatory Factor Analysis Results for Three-Factor Initial Evaluation: Comparison of Initial Evaluation Factor Structures

Model	Configuration	χ^2	<i>df</i>	χ^2/df	SRMR	RMSEA	CFI
Model 1	One-factor	263.02**	27	9.74	0.10	0.24	0.86
Model 2	Two-factor (Comp & Sim) Like	179.81**	26	6.92	0.10	0.20	0.91
Model 3	Two-factor (Comp & Like) Sim	151.50**	26	5.83	0.07	0.18	0.92
Model 4	Two-factor (Sim & Like) Comp	147.78**	26	5.68	0.08	0.17	0.93
Model 5	Three-factor	40.21*	24	1.68	0.05	0.06	0.99

Note. *N* = 189. SRMR = standardized root-mean-square residual; RMSEA = root-mean-square error of approximation; CFI = comparative fit index; Comp = competence; Sim = similarity; Like =liking.

p* < .05. *p* < .01.

(RMSEA) = .24; comparative fit index (CFI) = .86. The results reported in Table 1 also indicate that all three two-factor models had poor fit. Our three-factor model based on Howard and Ferris' (1996) theoretically distinct set of interviewer evaluations was the only model we tested that exhibited good model fit, Model 5, $\chi^2(24) = 40.21, p < .05; \chi^2/df = 1.68; SRMR = .05; RMSEA = .06; CFI = .99$. This provides support for three distinguishable initial evaluations.

Interview score. Structured questions and response scales were developed with subject matter experts' (SMEs') judgments of job relevance. Three SMEs who had been involved in interviews for accountants reviewed the interview questions and rating scales to establish content validity. All SMEs agreed that the interview assessed three constructs that were relevant to the job: interpersonal skills, problem solving, and work motivation. Inspection of sample performance evaluation forms used by all four accounting firms to assess a PPA student's internship performance also underscored the relative importance of these questions.

We developed 12 questions: six items to assess interpersonal skills, four items to assess problem solving, and two items to assess work motivation. In addition, the SMEs noted that the rating scales for each question enabled the interviewer to accurately evaluate the candidate's answers relative to the constructs. An example of an interview question is "Tell me about a time that you had to work with someone you did not like" (for interpersonal skills), "Tell me about a time when you had to make a quick decision with limited information" (for problem solving), and "Tell me about a time when you worked hard to obtain a goal despite difficult obstacles" (for work motivation). Interviewers evaluated responses using a 5-point response scale, ranging from *poor* (1) to *superior* (5), which was anchored by behavioral descriptions of each particular competency at each level (Campion et al., 1997). The interview score was calculated as the mean of the 12 items ($\alpha = .84$)

Proportion of internship offers. Each participant in Sample 1 was sent a follow-up survey approximately 1 month after all the interviews were conducted and internship offers had been extended by the Big Four accounting firms. Applicants reported the number of Big Four firms they interviewed with and internship offers they received. Although the Big Four accounting firms had to interview all PPA students, the students could choose the number of interviews in which to engage, ranging from none to four interviews with the Big Four accounting firms. We controlled for the number of firms that the candidate agreed to interview with in order to capture their internship offer success rate—the number

of Big Four internship offers adjusted for the number of firms with which students chose to interview. Because internship offers from the Big Four accounting firms were all extended on one specific day, students received internship offers from their "top choice" and "bottom choice" at the same time. Therefore, they could not "remove" themselves from consideration for other offers with any certainty that they had the offer they wanted. Students interviewed with an average of 3.89 of the Big Four firms and received an average of 2.15 offers.

Invitations to second interviews. To examine whether candidates' "exaggeration" of offers distorted relationships, we obtained responses from each of the Big Four accounting firms for Sample 2. Each was sent a follow-up request to provide names of those PPA applicants who were invited to a second job interview. Three of four firms provided names of those invited back for a second employment interview. For Sample 2, all PPA students interviewed with all Big Four firms. Thus, the sum of invitations was used as the score reflecting number of invitations for second interviews.

Control variable. Prior research on interviews has shown that non-job-related "biasing" effects often disappear after candidate qualifications are taken into consideration (Schmitt, 1976). The PPA coordinator stated, and human resources managers in the firms verified, that accounting firms relied on grade point average (GPA) in the accounting major courses as a critical indicator of qualifications for the job. Consequently, GPA served as a control for work qualifications.

Results

Table 2 presents means, standard deviations, and zero-order correlations among the variables combined from Samples 1 and 2 on the basis of just one interviewer's ratings. Although most of the correlations are consistent with expectations, one correlation warrants comment. The sum of the interviewer evaluations from the structured questions (interview score) asked during the mock interview significantly correlated with the proportion of internship offers extended later by the Big Four accounting firms, $r = .41; 95\%$ confidence interval (CI) [.22, .57]. This shows those candidates obtaining higher ratings during the structured interview from the mock interviewers also received a higher proportion of internship offers from the Big Four accounting firms.

Table 3 reports means, standard deviations, and zero-order correlations among the variables used for Sample 2 based on ratings from different interviewers. The pattern of correlations is

Table 2
Means, Standard Deviations, and Intercorrelations Among Study Variables From the Same Interviewer

Variable	M	SD	1	2	3	4	5	6	7	8	9
1. GPA	3.53	.33	—								
2. Sex	0.64	.48	.08	—							
3. Age	20.44	.66	-.07	.04	-.07						
4. IE overall rating	3.56	.61	.09	.01	-.17*	—					
5. IE liking	4.12	.67	.11	.15	-.12	.64*	—				
6. IE similarity	3.45	.63	.04	.09	-.02	.52*	.55*	—			
7. IE competence	4.17	.77	.11	.05	-.22*	.60*	.70*	.48*	—		
8. Interview score	3.53	.53	.23*	-.10	-.04	.42*	.34*	.39*	.42*	—	
9. Internship offers	0.54	.30	.29*	.10	.02	.22*	.24*	.24*	.30*	.41*	—

Note. N = 189, except for N = 76 with internship offers. Sex coded as 0 = male, 1 = female. GPA = grade point average; IE = initial evaluations. * p < .05.

similar to that reported in Table 2, but as expected, the relationships are weaker due to different interviewers providing ratings. For example, the correlation between the rating of competence during the initial impression and structured interview score was .42, 95% CI [.23, .58], when the same interviewer rated both constructs (Table 2). However, when different interviewers provided the structured interview score (Interviewers 1 and 3) rather than the initial impression competency rating (Interviewer 2), the correlation was smaller yet still significant, r = .25 and .24, respectively; lower 95% CI = .07; see Table 3. Thus, all of the correlations between ratings of competence assessed during rapport building and the later structured interview scores were significant.

Hypothesis 1 predicted that a general overall evaluation based on an initial impression in our mock interview would predict which applicants received a higher proportion of internship offers. This prediction was supported, r = .22; 95% CI [.01, .44] (N = 76), as shown in Table 2. Initial overall impressions formed during the rapport-building phase of a mock interview thus predicted an actual selection decision.

We next examined whether early impressions are based on perceptions of competence, beyond the influence attributed to liking and similarity to the candidate. We conducted hierarchical regression analyses, which are reported in Tables 4–7 for each outcome, to assess the incremental effect of competence ratings. In

order to minimize the potential effect from the candidate’s work qualifications, we controlled for applicant GPA in the first step of all regression analyses. Candidate GPA was positively related to structured interview scores from the same interviewer, β = .23, 95% CI [.09, .39] (Table 4) and different interviewers, β = .23, 95% CI [.05, .41], and β = .23; 95% CI [.05, .41], respectively, in Tables 6 and 7. Candidates with higher GPAs also received more invitations to second follow-up employment interviews with actual employers, β = .21, 95% CI [.05, .41] (Table 5).

In the second step, the effects from ratings of early impressions of liking and similarity to the candidate were added to GPA. Only in Table 4 was either rating of the initial impression of liking for or similarity to the candidate meaningfully related to the dependent variable: β = .16, 95% CI [.01, .31], and β = .29, 95% CI [.14, .45], respectively. Of the four analyses, this was the only one in which the same interviewer rated both the early impressions of the candidate and the candidate’s answers to the structured interview questions; in Tables 5, 6, and 7, a different interviewer rated the interview outcome (whether performance in an interview or who to invite back for a follow-up interview). Thus, in all other analyses, the evaluation of the candidate’s initial impression, following rapport building, was obtained in a different interview conducted by a different interviewer. As a whole, these results suggest that when different interviewers provided the dependent variable, nei-

Table 3
Means, Standard Deviations, and Intercorrelations Among Study 2 Variables From Different Interviewers

Variable	M	SD	1	2	3	4	5	6	7	8	9
1. GPA	3.62	0.26	—								
2. Sex	0.61	0.54	.06	—							
3. Age	20.50	0.69	-.09	-.16*	—						
4. IE liking (in Interview 2)	4.38	0.83	.18†	.12	.04	—					
5. IE similarity (in Interview 2)	3.74	0.80	-.03	.11	.19	.61*	—				
6. IE competence (in Interview 2)	4.42	0.79	.17†	.23*	-.02	.74*	.49*	—			
7. Interview score (full; in Interview 1)	3.55	0.54	.21*	-.10	-.04	.24*	.10	.25*	—		
8. Interview score (structured only; in Interview 3)	3.55	0.58	.23*	.13	-.21*	.13	.03	.24*	.31*	—	
9. Second interview	1.58	1.00	.23*	-.08	.02	.19*	.16†	.23*	.22*	.18†	—

Note. N = 113. Sex coded as 0 = male, 1 = female. Interview 2 is a rapport-building interview only; Interview 1 is a full interview with rapport building and structured interview; Interview 3 is a structured interview only; Second interview = sum of times invited to a second interview from three of the Big Four firms. GPA = grade point average; IE = initial evaluations. † p < .10. * p < .05.

Table 4

Regression of Structured Interview Score on Initial Evaluations of Competence, Liking, and Similarity (Combined Samples 1 and 2) With the Same Interviewer

Independent variable	b	se _b	β	R ²	Adjusted R ²	ΔR ²	dfs
Step 1				.05	.05	—	1, 187
Applicant GPA	0.39**	0.12	.23**				
Step 2				.21	.20	.16**	3, 185
Applicant GPA	0.34**	0.11	.20**				
IE liking	0.12*	0.06	.16*				
IE similarity	0.25**	0.07	.29**				
Step 3				.26	.24	.05**	4, 184
Applicant GPA	0.32**	0.11	.19**				
IE liking	−0.02	0.08	−.03				
IE similarity	0.22**	0.07	.25**				
IE competence	0.21**	0.06	.30**				

Note. N = 189. IE = initial evaluations; GPA = grade point average.
* p < .05. ** p < .01.

ther the effect from initial liking nor the effect from similarity to the first interviewer generalized.

In the third and final analysis step, we examined whether the early impressions of candidate competence during a structured interview predicts interview outcomes above and beyond the effects due to liking for and similarity to the candidate, whether from the same interviewer (i.e., Hypothesis 2, in Table 4), from different interviewers (Hypotheses 4 and 5, Tables 6 and 7), or from those making actual selection decisions (Hypothesis 3, Table 5). Results of these regressions yielded consistent support for the hypothesized effect for evaluations of candidate competence above and beyond the effects from liking and similarity— $\beta = .30$, 95% CI [.13, .48]; $\beta = .35$, 95% CI [.08, .62]; and $\beta = .33$, 95% CI [.05, .60], respectively—for predictions of structured interview scores by the same (Table 4) or different interviewers (Tables 6 and 7). These results provide support for Hypothesis 2, which posited competence evaluations have an incremental effect beyond liking and similarity. This effect persisted when structured interview ratings were from a different interviewer, supporting Hypothesis 4, and from still another interviewer who did not even engage in rapport-building activities, supporting Hypothesis 5. Hypothesis 3 was not supported, as the incremental effect was not significant

after we accounted for applicant GPA and ratings of liking and similarity, $\beta = .17$, 95% CI [−.11, .44] (Table 5). However, the zero-order correlation between interviewer evaluations of candidate competence made during rapport building was significantly related to more invitations to follow-up employment interviews with actual employers, $r = .23$, 95% CI [.06, .41].

Discussion

Our results reveal that an initial impression formed during the “unstructured” rapport-building segment at the beginning of the interview is indeed related to interviewer evaluations during the structured interview and even predict later employment decisions. Our findings also show that interviewer ratings of candidate competence during rapport building accounted for incremental variance in outcomes beyond the variance attributed to perceived similarity or initial likeability of the candidate. Results also show that initial evaluations of competence made in one (mock) interview predicted the proportion of internship job offers from a number of different firms (and interviews). These findings contradict assumptions that the only information that matters in the interview is that conveyed during a series of job-related questions.

Table 5

Regression Interview Outcomes (Invitations to a Second Interview From Actual Employers) on Initial Evaluations of Competence, Liking, and Similarity (Sample 2) With Different Interviewers

Independent variable	b	se _b	β	R ²	Adjusted R ²	ΔR ²	dfs
Step 1				.05	.04	—	1, 111
Applicant GPA	0.89*	0.36	.21*				
Step 2				.09	.06	.03	3, 109
Applicant GPA	0.85*	0.37	.22*				
IE liking	0.09	0.14	.07				
IE similarity	0.16	0.15	.13				
Step 3				.10	.06	.01	4, 108
Applicant GPA	0.82*	0.37	.21				
IE liking	−0.05	0.18	−.04				
IE similarity	0.14	0.15	.11				
IE competence	0.21	0.17	.17				

Note. N = 113. Initial evaluation (IE) ratings from Interview 2. GPA = grade point average.
* p < .05. ** p < .01.

Table 6
Regression of Structured Interview Score on IE Competence, IE Liking, and IE Similarity (Sample 2) With Different Interviewers

Independent variable	b	se _b	β	R ²	Adjusted R ²	ΔR ²	dfs
Step 1				.05	.04	—	1, 111
Applicant GPA	0.67*	0.28	.23*				
Step 2				.08	.05	.03	3, 109
Applicant GPA	0.57*	0.28	.19*				
IE liking	0.17	0.11	.19				
IE similarity	−0.05	0.11	−.05				
Step 3				.13	.10	.05*	4, 108
Applicant GPA	0.52	0.28	.18				
IE liking	−0.05	0.14	−.06				
IE similarity	−0.07	0.11	−.07				
IE competence	0.34*	0.13	.35*				

Note. *N* = 113. Initial evaluation (IE) ratings from Interview 2 and structured interview score in Interview 1. GPA = grade point average.
* *p* < .05. ** *p* < .01.

Moreover, our results show the “fast and frugal” judgments formed during the meet-and-greet phase of the interview incorporate, in part, subjective ratings of the candidate’s competence to do the job.

Theoretically, our results highlight the need to incorporate the use of a two-stage decision model in research about the structured interview (Dipboye, 2004; Stevens, 2000). Researchers must further consider how fast, automatic, intuitive information exchanged during rapport building affects interviewers’ decisions. In many respects, the critical issue is whether these initial impressions are driven by job-relevant antecedents. We would expect initial evaluations of competence to have higher validity if they are predicted by job-relevant candidate characteristics such as general mental ability, emotional stability, or conscientiousness rather than job-irrelevant characteristics such as whether the candidate smiles or nods enthusiastically during the interview. These issues should be investigated in future research into the causes of initial evaluations in the interview setting.

In practice, the usefulness of initial impression ratings may depend on when they are collected during the selection process. For instance, initial evaluations of competence made early in the selection process, such as at a career fair or during a screening interview, where the selection ratio is expected to be quite low,

could provide useful information to help screen out obviously unqualified candidates. When selection ratios are low (most candidates are rejected), even predictors with modest validity can substantially improve utility (Taylor & Russell, 1939). Conversely, it would likely be unwise to base a final hiring decision on initial evaluations. Essentially, we suggest that initial evaluations might be suitable as a select-out, rather than a select-in, decision aid for some jobs (Kuncel, 2008).

Practically, a second way initial impressions of candidate competence may prove useful is when one is hiring for “relevant” jobs. Jobs requiring a high degree of interaction with the public or dealing directly with customers may be particularly relevant. For example, some casinos rate the candidate’s ability to interact with a stranger (i.e., interviewer) as an important part of the hiring recommendation (Gatewood, Feild, & Barrick, in press). These firms assume that how efficacious the candidates are at interacting during the rapport-building stage of an interview indicates how socially competent these candidates will be with customers during brief encounters later on the job.

Finally, although some may argue that eliminating rapport-building sessions from the interview may be an effective way of minimizing the effect of initial evaluations, we caution against this for two reasons. First, our results indicate that even when rapport

Table 7
Regression of Structured Interview Score (With No Rapport Building) on Initial Evaluations of Competence, Liking, and Similarity (Sample 2) With Different Interviewers

Independent variable	b	se _b	β	R ²	Adjusted R ²	ΔR ²	dfs
Step 1				.05	.04	—	1, 111
Applicant GPA	0.51*	0.21	.23*				
Step 2				.06	.03	.01	3, 109
Applicant GPA	0.46*	0.22	.20*				
IE liking	0.08	0.08	.11				
IE similarity	−0.03	0.09	−.04				
Step 3				.11	.07	.05*	4, 108
Applicant GPA	0.43*	0.21	.19*				
IE liking	−0.08	0.11	−.11				
IE similarity	−0.05	0.09	−.06				
IE competence	0.24*	0.10	.33*				

Note. *N* = 113. Initial evaluation (IE) ratings in Interview 2 and structured interview score in Interview 3.
* *p* < .05. ** *p* < .01.

building is eliminated from the interview, the structured interview score is still related to initial evaluations of competence (as rated by another interviewer). Second, rapport building is positively related to interviewer affective reactions to the interview while also functioning as an opportunity to recruit candidates (Chapman & Zweig, 2005). Thus, it seems premature to conclude that rapport building should be eliminated. Instead, we call for more research to investigate the validity of these initial impressions, especially in jobs in which employees must be perceived as competent following brief social exchanges with customers (e.g., restaurant servers, salespeople, hospital nurses).

Like any study, the current study has limitations. First, use of mock interviews may limit the generalizability of our findings to actual employment interviews (Posthuma et al., 2002). However, several elements of our study design minimize this threat to generalizability, including our use of motivated interviewers and candidates, two-way social interactions (rather than “paper people”), and actual outcomes from employment interviews conducted shortly after the mock interviews. Hence, fidelity in this study has been maximized to the extent possible. Moreover, our ratings predicted real-world outcomes. Another limitation is the modest sample sizes of some analyses ($N = 76$ in one case), as greater statistical power is desirable. Another is the lack of interview experience for some of the interviewers in this study. While researchers have suggested various ways that interviewer experience affects interview outcomes, the scarcity of research in this area has made a definitive conclusion premature (Posthuma et al., 2002). However, the extent that experience influences interviewer ratings in this study appears to be minimal, as we could detect no mean score differences for those with and without prior interview experience.

In conclusion, the interview at its core is an agenda-driven social exchange between strangers. One important objective for the interviewer is to judge the candidate’s qualifications for the job. In contrast, the candidate’s primary objective is to get a job. These conflicting agendas, compounded by the complex social dynamics of meeting a stranger, illustrate the challenges inherent in obtaining accurate information about candidate suitability during the interview. While considerable attention has been given to the ways in which the structure of the employment interview affects predictive validity (Jelf, 1999), researchers have failed to adequately consider how the meet-and-greet phase impacts interview decisions (Jelf, 1999; Posthuma et al., 2002).

Our study is the first to empirically show a link between interviewers’ initial evaluations of candidates’ competence, liking, and similarity formed during the rapport-building stage of the structured interview and a number of interview outcomes. These results reveal that initial interviewer impressions made on the basis of limited information, particularly those reflecting competence, predict interviewer ratings of candidate responses to structured questions, as well as actual employment decisions including whether the candidate was invited back for a second interview or received an employment offer. This implies the “fast and frugal” judgments that emerge at the start of an interview matter when interviewers are making a selection recommendation. In the future, researchers must examine which factors (e.g., professional appearance, impression management, and so forth) affect these initial impressions and to what extent these initial impressions predict job performance.

References

- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*, 256–274.
- Anderson, C. W. (1960). The relation between speaking times and decision in the employment interview. *Journal of Applied Psychology*, *44*, 267–268.
- Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, *6*, 269–278.
- Barrick, M. R., Shaffer, J., & DeGrassi, S. D. (2009). What you see may not be what you get: Relationships among self-presentation tactics and ratings of interview and job performance. *Journal of Applied Psychology*, *94*, 1394–1412.
- Buckley, M. R., & Eder, R. W. (1988). B. M. Springbett and the notion of the “snap decision” in the interview. *Journal of Management*, *14*, 59–69.
- Cable, D. M., & Judge, T. A. (1997). Interviewers’ perceptions of person–organization fit and organizational selection decisions. *Journal of Applied Psychology*, *82*, 546–561.
- Campion, M. A., Palmer, D. K., & Campion, J. E. (1997). A review of structure in the selection interview. *Personnel Psychology*, *50*, 655–703.
- Chapman, D. S., & Zweig, D. I. (2005). Developing a nomological network for interview structure: Antecedents and consequences of the structured selection interview. *Personnel Psychology*, *58*, 673–702.
- Dipboye, R. L. (2004). The selection/recruitment interview: Core processes and contexts. In A. Evers, N. Anderson, & O. Voskuilj (Eds.), *The Blackwell handbook of personnel selection*. Oxford, England: Blackwell.
- Funder, D. C., & Colvin, R. C. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology*, *60*, 773–794.
- Gatewood, R. D., Feild, H. S., & Barrick, M. R. (in press). *Human resource selection* (7th ed.). Mason, OH: Thompson South-Western.
- Higgins, C. A., & Judge, T. A. (2004). The effect of applicant influence tactics on recruiter perceptions of fit and hiring recommendations: A field study. *Journal of Applied Psychology*, *89*, 622–632.
- Howard, J. L., & Ferris, G. R. (1996). The employment interview context: Social and situational influences on interview decisions. *Journal of Applied Social Psychology*, *26*, 112–136.
- Huffcutt, A. I., & Arthur, W. (1994). Hunter and Hunter (1984). revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology*, *79*, 184–191.
- Jelf, G. (1999). A narrative review of post-1989 interview research. *Journal of Business and Psychology*, *14*, 25–58.
- Kahneman, D. (2003). A perspective on judgment and choice. *American Psychologist*, *58*, 697–720.
- Kuncel, N. R. (2008). Some new (and old) suggestions for improving personnel selection. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *1*, 343–346.
- Posthuma, R. A., Morgeson, F. P., & Campion, M. A. (2002). Beyond employment interview validity: A comprehensive narrative review of recent research and trends over time. *Personnel Psychology*, *55*, 1–81.
- Schmitt, N. S. (1976). Social and situational determinants of interview decisions: Implications for the employment interview. *Personnel Psychology*, *29*, 79–101.
- Springbett, B. M. (1958). Factors affecting the final decision in the employment interview. *Canadian Journal of Psychology*, *12*, 13–22.
- Stevens, C. K. (2000). Structure interviews to hire the best people. In E. A. Locke (Ed.), *Handbook of principles of organizational behavior* (pp. 29–40). Malden, MA: Blackwell.
- Stevens, C. K., & Kristof, A. L. (1995). Making the right impression: A field study of applicant impression management during job interviews. *Journal of Applied Psychology*, *80*, 587–606.
- Stewart, G. L., Dustin, S. L., Barrick, M. R., & Darnold, T. C. (2008). Exploring the handshake in employment interviews. *Journal of Applied Psychology*, *93*, 1139–1146.

- Taylor, H. C., & Russell, J. T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection: Discussion and tables. *Journal of Applied Psychology, 23*, 565–578.
- Van Iddekinge, C. H., Raymark, P. H., & Roth, P. L. (2005). Assessing personality with a structured employment interview: Construct-related validity and susceptibility to response inflation. *Journal of Applied Psychology, 90*, 536–552.
- Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology, 57*, 120–128.
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*, 592–598.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151–175.

Received December 3, 2008

Revision received March 30, 2010

Accepted April 9, 2010 ■

Members of Underrepresented Groups: Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write APA Journals at Reviewers@apa.org. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.
- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.
- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In the letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, “social psychology” is not sufficient—you would need to specify “social cognition” or “attitude change” as well.
- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.