

Interviewer Cheating: Implications for Research on Entrepreneurship in Africa

David E. Harrison and Stefanie I. Krauss

Abstract

Interviewer cheating has seldom been studied or discussed as a problem in the literature. This article therefore begins with a brief review of this problem area, which is of utmost importance especially for entrepreneurship research in a Third World context. In the Third World, the allocation of financial support is often based on interview surveys. After describing two cases of extensive faking by carefully selected and comprehensively trained interviewers, possible explanations of such behavior are hypothesized. The paper concludes with the warning that interviewer cheating may be more prevalent than is generally assumed. It is recommended that careful interviewer selection, preventive study design and meticulous subsequent interview data inspection should be used whenever possible.

Key words: Interview cheating, interview faking, fake data, data reliability

In spite of early evidence of cheating in market research and other fields where professional interviewers are employed (e.g., Case, 1971; Durant, 1946; Evans, 1961; Harrison, 1947), the problem of interviewer cheating has largely been ignored in recent literature. In preparing the current article, the authors carried out a detailed literature search on CD-ROM databases in the domains of psychology, economics, marketing, and the social sciences. Search criteria utilized were 'interview cheating', 'interview faking', 'fake data', and 'data reliability'. This search revealed little relevant research in this area. Recognized as early as the 1940s (e.g., Crespi, 1945; Durant, 1946; Elinson & Cisin, 1948; Harrison, 1947), the cheating problem has not, except for a brief renaissance in the 1980/90s (e.g., Biemer &

David E. Harrison is a psychologist and is currently developing the psychological, industrial and commercial consultancy, training and research activities of Human Resources (Pvt.) in Zimbabwe. Stefanie I. Krauss is a psychologist and is a research associate and a PhD student at the Department of Work and Organizational Psychology at the University of Giessen in Germany. The authors gratefully acknowledge the funding of this research through the Deutsche Forschungsgemeinschaft (German Research Community, project no.: FR 638/13-1).

Stokes, 1989; Kiecker & Nelson, 1996), been subsequently investigated. This article aims to review the existing literature, create problem awareness, re-initiate a discussion in the entrepreneurship research community that greatly relies on interview-based data, and suggests some solutions for the enhancement of data accuracy, reliability, and validity.

Background

The difficulty of assessing levels of cheating and the scarcity of definitive research on cheating are factors that may have contributed to the brevity of commentary on this aspect of research in major texts. For example, the Kinnear & Taylor (1979) textbook contains only seven lines on interview cheating and asks the question: What is cheating? In practice, it is hard to define. Obviously, the personal interviewer who sits at home and fills out the questionnaire is cheating. More frequently, cheating is defined as the falsification of a question or questions within the questionnaire. This type of cheating is extremely difficult to detect" (p. 480). Besides this most severe type of interviewer fraud, there are other types that are even harder to detect and assess such as the appropriateness of the interview respondent as specified in the study design and appropriateness of the interviewing procedure (Case, 1971).

Some researchers tried to control interviewer cheating with subsequent inspection for incongruities, for repetition of answers, for overly meticulous ballots, and for obvious revisions (Crespi, 1945). It is on this follow-up that the quality of the interview is then assessed. Additionally, a few studies included follow-ups on the correct interview procedures mostly with postcards, which the interview partners are asked to send back to the research organization. Stewart and Flowerman (1951) reported the use of concealed wire recorders to check on cheating and error behavior of fifteen public opinion interviewers in thirty-three recorded interviews. 'Traditional' quality rating procedures were compared with criterion cheating scores derived from the recorded interviews. The authors were disconcerted by their findings:

It is regrettable but necessary to report, that the correlation between these quality ratings and the criterion cheating scores was $-.1$; that is, interviewers who committed a greater amount cheating received somewhat better quality ratings than those who had cheated less. The correlation between the quality ratings and the criterion error scores was $.18$, a correlation too low to afford proof of any validity for the quality ratings. We were forced to conclude that if our data were representative, the traditional quality ratings were either useless or worse than useless as predictors of cheating or error behavior (p. 166–167).

Similarly, the method of using postal check-ups on interviewers, which are commonly employed to avoid the high costs of personal quality control, are not very reliable. Return rates of mail surveys are generally not very high and in the case of the detection of interviewer cheating, it is to be expected that precisely those who have not been interviewed will fail to return any information to the researchers. This implies that physical checking of the interviewer behavior is indispensable—especially because interviewers often provide accurate demographic and factual data with a shortened version of the actual interview while everything else stems from their imagination (Evans, 1961).

Surprisingly, despite the massive growth in interviewer dependent research since the Stewart and Flowerman (1951) findings, there has been minimal systematic research against precautionary criteria as strong as concealed recording of interviews. Case (1971) cautioned that one out of four survey interviews might contain serious errors. He found that, despite using an interview verification service (FACT-Field Audit and Completion Test), which was developed in 1968 by the Advertising Research Foundation and with the interviewers being aware of this checking procedure, 17.6% of interviews had frequent performance errors and/or non-verified interviews and only 46.6% had no errors at all. It should be noted that there was a strong correlation (76%) between interviewers with non-verified interviews and interviews with performance errors. In spite of these discouraging results, it is difficult to find subsequent studies in which interviewer misbehavior has been cited as a significant problem.

This scarcity of follow-up research might have to do with the difficulties in the detection of interviewer cheating. Other reasons include the possibly negative implications of such findings for interview-based research and market research. Profit making organizations (which includes universities to some extent) are certainly not interested in publications on interview fabrications, particularly regarding their own research. Many cases of cheating could have been prevented through a more careful study design and may therefore be ascribed to the researchers' own abilities. Crespi (1945) described cheating more as a problem of interviewer morale than as a problem of interviewer morals and identified ballot and administrative demoralizers. As ballot demoralizers that might make interviewing more difficult and hence lead to cheating, he categorized unreasonable length (of the interview or questionnaire), too many why and what for questions, apparent repetition of questions, lengthy wording of questions, complex and difficult questions, and antagonizing questions. As administrative demoralizers, he categorized the interviewer as having no personal but mainly mail contact with the research organization, overly difficult sample assignments, external factors (such as the weather or bad roads), and part-time work (where other jobs and assignments might interfere with interviewer's thoroughness).

Somewhat contradicting Crespi (1945); Durant (1946) argues that the very nature of full-time interviewing is demoralizing because it is not a full-time job and

“one day’s interviewing, however well done, merely serves to lead on to the next day’s interviewing”. Additionally, he argues that “the work is not, in itself constructive” and that “there is probably something personally disintegrating about the process of continually taking people’s views and habits to pieces” (p. 290).

Even though called for at an early stage of the cheating debate (Durant, 1948), there has been no attempt to identify interviewer personality. The only constructive step in this direction is probably the notion by Crespi (1946) that interviewers should not score too high on intelligence because especially highly intelligent interviewers are more likely to be demoralized by the often adverse job conditions. Additionally, only rather intelligent interviewers could fabricate interviews in such a way that they are hard to detect. However, Crespi noted that the downside of hiring less intelligent interviewers would be to sacrifice competence and skill in favor of morale.

The above literature overview clearly shows that the debate on interviewer misbehavior currently seems to have been halted and that there are still many open questions. To re-start the discussion, the experiences with interview fabrication will be reviewed and further conclusions will then be drawn on preventative study design, interviewer selection, and data inspection. This discussion is especially important in the African Third World context, where donor organizations as well as governmental and non-governmental research organizations make high investments in interview surveys, which are often referred to in policy formulation and for the allocation of financial support and project funding. Also, the in-depth and extended interviews necessary for research into entrepreneurship in the African context place heavy burdens on both the interviewers and their trainers. In such a context, commitment by the interviewer may need to be much greater than in more straightforward enumerative studies.

Two Cases Of Multiple Interview Fabrications

A research group has been conducting studies on entrepreneurship in the African context for several years. While former data sets were completed with the help of M.A. students who were interested in the data for their own final theses, the most recent Zimbabwean study was carried out with hired interviewers. This was necessary because of a desire to collect a relatively large sample (320 interviews) in rural areas where English is less frequently spoken. After a pilot study in 1997, interviewing was started for measurement time one in 1998/99. Two hundred out of the 320 interviews were conducted by a local logistics company.

The first of the three interview fabricators whose inexplicable behavior destroyed much of our data and stimulated this paper was Daniel (assumed name). As was true with the other interviewers, Daniel was selected using the logistics company’s middle level selection battery consisting of a total of ten locally normed and validated tests of personality characteristics, occupational interests, aptitudes,

and intelligence. His overall test performance was at least equivalent to the graduates that followed him, although he had only four years of secondary education. His task as interviewer was to administer business-related questionnaires, with both focused and open-ended questions, tests and exercises over a period of approximately 2½ hours. As it was necessary to motivate the interviewee to respond to questions and tests which challenged memory and cognitive processes for a sustained gap in the normal work day, interviewees were paid Zimbabwe\$ 300.00 per interview (which at the time equaled five US\$).

Daniel worked in the field with researchers from the University of Giessen, Germany. He and the other interviewers sometimes shared lodging and spent much time together. The logistics manager and the researchers checked his data and went out in the field with him for supervisory feedback-interviews on his interviewing technique (which was assessed as excellent by all parties involved). However, his interviews were not checked with random repeat visits. It was undeniably an error to leave out this standard procedure, but re-visits were logistically difficult and appeared unnecessary, such was our general confidence in Daniel's professionalism.

The second part of this longitudinal study took place about one year later. In order to retain his services for the second series of interviews, the logistics company employed Daniel during the interim period and gradually increased his salary to 300% of the original. In the second series of interviews he worked even more closely with the visiting researchers from Germany. However, a further error was committed on the part of the researchers, as Daniel had access to the protocols of the first series of interviews. Accordingly, inconsistencies and contradictions that would have been inevitable had he been required to complete the second 'faking' from memory, did not show up as he only needed to make minor changes in fabricating the second interviews.

After approximately five weeks, one of the researchers from Germany nevertheless spotted an anomaly that led to the investigation of Daniel's interviews. After finding at least 15 cases which Daniel had faked on both the first and second occasions, we questioned him and he eventually admitted having fabricated interviews. He was unwilling to state how many. After a thorough follow-up on all of his interviews, it transpired that he had faked all, except those when he was being trained, or when he was under direct observation—which totaled only 12% out of 148 of the time one series of interviews. Daniel was reported to the police for fraud and spent three months in jail followed by three months of community service.

Bearing this experience in mind, the interviews performed by a second interviewer, Gary (assumed name), at measurement time one were also checked. The investigation into Gary's interviewing behavior led to a similarly discouraging result. As the German research team had to leave the country immediately after Gary's interviewer training was finished, no direct observation on his interviewing was possible. This led to the unfortunate result that we found all of his 52

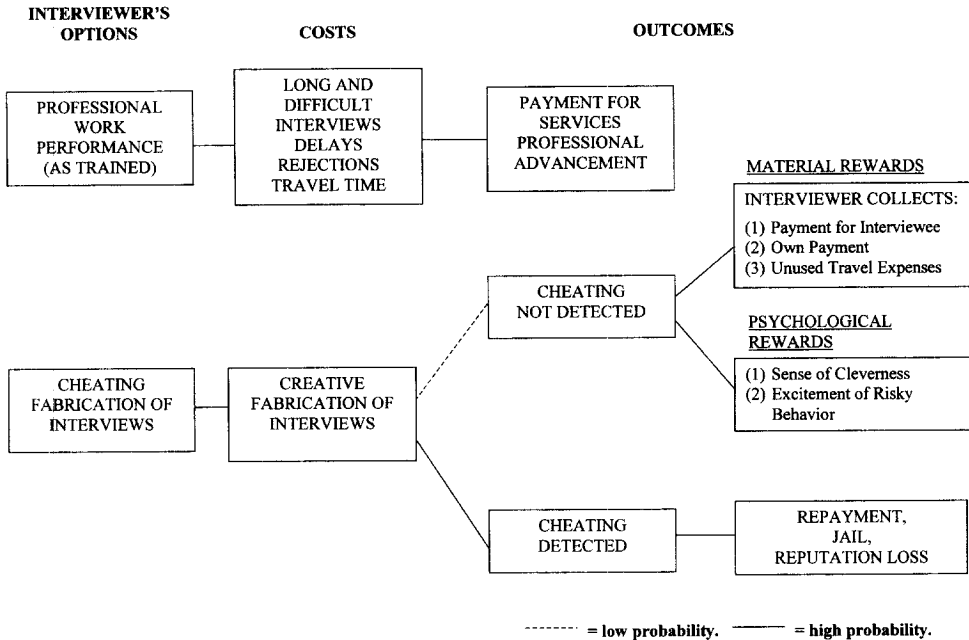
interviews to be fabricated. Regrettably, legal action could not be taken against Gary because his physical address had changed by the time his cheating was detected and finding him was therefore deemed impossible. Gary had been employed several months after Daniel and it was thought that Daniel had introduced Gary to his techniques of cheating. Daniel had interviewing experience and had already worked for the researchers in the pilot study, while Gary was a first-time interviewer. Newly trained interviewers are unlikely to fabricate interviews from the beginning, because they are not opportunistically prepared through knowledge of 'holes' in the system.

To summarize these experiences for measurement time one and the first part of measurement time two, it can be concluded that both interviewers, who did not have a personal interest in the accuracy of the data, cheated extensively. Consequently, the faked data was scrapped, and the researchers set about recruiting a new team of interviewers in order to fulfill the research obligation. A total of 237 replies to an advertisement were received, from which 65 respondents were identified and screened. From these, a short-list of 18 was selected for administration of the middle level selection battery. Five were then formally selected for training, which was carried out by a German researcher. During a week's training course the new interviewers became familiar with the interview and its rating scheme. In the first phase of training the interviewers practiced in a role-play setting. Thereafter, in vivo practices were carried out: The interviewers went into the field and performed approximately five training interviews in the presence of the trainer who subsequently gave extensive feedback.

As in the first interview series, such feedback sessions were also conducted at random intervals throughout the whole interview phase of the project. Thus, the training process was an ongoing one. Part of the briefing of interviewers appears below, as an illustration of the researcher's attempts to both caution and motivate: "This is your first job, for which you have been well trained. Success in this job will contribute towards your work experience and your curriculum vitae (resume). Please carry out your interviews as you have been trained to do, and you will not only receive a good reference from the survey organizers but you will also be eligible for a bonus. Your work will be checked by those who receive it, and there will be random checks on a sample of those whom you have interviewed. It is greatly in your interests to do a good job."

All five selected trainees for the interview series two were graduates from the University of Zimbabwe. One dropped out during training, and one during the first month of interviewing. Tony performed inadequately, and was retrained. This left the researchers with Carl, George and the retrained Tony (assumed names). It was decided to allow for no mistakes this time, with tape recordings demanded of all interviews, and the assignment of an a Shona (the local indigenous language) speaking member of staff to listen to all tapes. Interviewing proceeded apparently successfully for three months. However, while investigating an anomaly discovered

Figure 1
Utility Process Diagram of The Interviewer's Dilemma.



by our data transcriber, it was found that George had been declaring approximately one in three respondents as refusing to be taped. This claim was checked and it was discovered that all such refusals were fabricated interviews. George was reported to the police. He made partial restitution of the stolen interviewee fees, so charges were withdrawn.

At this stage, the self-destructive behavior of the interviewers was becoming as interesting as the data they were intended to collect. Of the five interviewers working for the researchers (Daniel and Gary in time one, Daniel, George, Tony, and Carl in time two) only two (Tony and Carl) produced reliable data and proved to be trustworthy. Three of the interviewers committed the most severe type of cheating behavior – data fabrication.

Possible Explanations

Applying utility theory to the interviewer's situation, a process diagram is presented below showing options and outcomes of interviewer behavior and in this study (see Figure 1). From this process diagram, it would appear that the professional (i.e., no cheating) route would be the obvious choice. However, three consecutive highly selected, comprehensively trained and able interviewers consistently faked data and stole funds intended for travel and for interviewee com-

pensation. The third even cheated despite having been told that his predecessor was in jail.

In the interests of future research, an explanation is needed as to why the utility process diagram is deficient in predicting the behavior of our interviewers. The following are possible explanations:

Cheating is Normative

When expressed in the literature, most cautions about interviewer cheating have been based upon the implicit assumption that honest interviewing is normative and that cheating is deviant. However, Kiecker and Nelson (1996, p.161) who studied telephone interviews concluded: misbehavior by telephone interviewers is not so much unusual and deviant as it “is ordinary and normal”. If this were so, then it would be reasonable to expect, and not be surprised by, a high level of cheating in the attritional environment of research in Africa. Perhaps the normative attitude is go for the sure, immediate pay-off and don’t hold too much faith in long-term opportunities or promises. In Zimbabwe, an impoverished country, most of the young are used to satisfying at the first opportunity, i.e. eat when food is available and in general gratify oneself whenever possible and do not miss opportunities. In our case the 300 Zimbabwe dollars, designed for the interview participant plus a travel allowance and payment for the faked interview, were large pay-offs, with the first two being immediate. This was a tempting situation, particularly in the context of the interviewer’s work among small business owners who were mainly operating in the informal sector (not registered or paying tax) and many of who were breaking local or government regulations.

Cheating is Self-Actualizing

Apparent success in fabricating an interview, being paid for doing so, and out-smarting the researcher may be an exhilarating experience, especially when shared with members of one’s own reference group.

Cheating is the Fault of the Researcher

In explaining interviewer cheating, some authors have tended to be intropunitive. For example, Bennett (1948) felt that interviewer misbehavior seldom arises until the interviewer encounters situations that are unexplained, difficult to understand, impractical, poorly managed, or unprofitable for the interviewer. By implication, get those things right and interviewer cheating will be minimal!

It has been suggested by Roth (1966) and others, that interviewers with a hired-hand mentality may be more prone to cheating. It would appear realistic to expect just such an attitude in the African research context, wherein the financier of the research is usually external to the culture in many respects, and even the research manager often differs from interviewers in terms of race, country of origin, education level, etc. The trainer in such cases tends to assume that the elegance and

importance of the research design will not be understood, and thereby depends upon using and emphasizing lists of ‘dos’ and ‘don’ts’. The trainer or survey manager is in a complex role relationship with trainee interviewers. On the one hand he/she must ‘test’ questions with the trainees to ensure cross-language identity of meaning. In such a role the trainee acts as the consultant. But, in carrying out interviews, the now-trained interviewer must behave with unquestioning obedience. In such circumstances ‘corner-cutting’ may begin with the occasional assisted response and, in difficult circumstances, may extend to completely faked interviews.

The Interviewers’ Explanations

Unfortunately, many hours of de-briefing Daniel and George have produced explanations by these interviewers, which leave the question unanswered. Daniel when asked why, in the light of benefits that would have accrued had he behaved professionally, did he fake the protocols, replied: “these things happen”. This was a vacuous comment from a person of Daniel’s obvious high ability and contributed minimally to our understanding of his sustained bout of cheating.

In a de-briefing before being handed over to the police, George was asked, “Why, when you were aware that the previous cheating interviewer, Daniel, had been sent to jail, when you were aware of the certainty that your work would be checked and faking discovered, did you persist in fabricating interviews?” He replied, “I’m sorry, I did fake some interviews but I hoped to make up for it by doing some good work later on” (at a higher level of insight, he might of course, have concluded that having thus spoiled the data, there would be little to be gained from adding better data at a later stage).

When we asked Carl (the remaining interviewer, whose work has been 100% checked as exemplary) why George might have cheated despite Daniel having been sent to jail, Carl replied: “We were not really convinced that an interviewer had been jailed. Also, George intends to be an accountant so he wasn’t as motivated as I am. I intend to work in this area.”

Conclusions

Clearly, more research is needed, in a variety of contexts and with interviewers from several populations before we will be able to provide satisfactory and practically useful predictions and explanations of interviewer cheating. It could be suggested that the experiences described here have been the outcomes of bad luck and weak management. However, the literature review has been augmented with informal discussions held with several interviewers, in order to find out more about the incidence of interviewer misbehavior. At this anecdotal level, many cases have been noted that are similar to those reported in this paper. Derived from these various sources of information, a catalogue of interviewing procedure rules have

been developed that could help researchers to improve considerably the accuracy, reliability, and validity of their data:

Study Design

- Ensuring reasonable interview length, avoiding apparent repetitions of questions, lengthy wording, and complex, difficult, and antagonizing questions (Crespi, 1945).
- Creating positive work conditions and avoiding unrealistic production quotas, time constraints, or work rules (Kiecker & Nelson, 1996).
- Keeping a close contact between interviewers and researchers and avoiding mail contact only (Crespi, 1945).
- Tape recording all interviews (if money is paid, the transaction should also be on tape).
- Insisting on receipts with signature/letter of acknowledgement with signature of all participants, especially if money is paid for participation.
- Promotion of group norms and corporate cultures among interviewers and their supervisors that emphasize data accuracy (Kiecker & Nelson, 1996).
- Interviewer compensations on a per hour, not a per interview basis (Kiecker & Nelson, 1996).
- Use of non-existing addresses/phone numbers in the participants' lists to detect interviewer cheating (Kiecker & Nelson, 1996).

Interviewer Selection

- Employment of interviewers with personal interest in the data quality (e.g., students).
- Little interviewer experience to reduce likelihood of knowledge on cheating opportunities in the system (this has to be offset against potential advantages of interviewer experiences)
- Measurement of interviewer characteristics such as morals, morale, hired-hand attitudes, level of boredom (Kiecker & Nelson, 1996) and the employment of integrity testing for interviewer selection.

Subsequent Interview Inspection

- Data inspection hints: Comparison of handwriting and ticking habits on questionnaires, and stereotypical answer patterns between participants (should differ between subjects).
- Data inspection for within participant hints: Incongruities, repetition of answers, overly meticulous interviews, and revisions (Crespi, 1945).
- Inspection of factual interview data (e.g., do addresses or phone numbers given by the interviewer exist?).
- Interview verification: Random checks on a subsample of participants with a short questionnaire (was the interview carried out at all and was it carried

out correctly and according to the research design (Biemer & Stokes, 1989). Important: Fieldwork should be sent in daily and be verified at once (Evans, 1961). Verification should best be done by an outside organization to avoid leaking back of information to the interviewers (Case, 1971).

Additional Inspection Issues for Longitudinal Studies

- Comparison of signature and handwriting at measurement times one and two.
- Inspection of participants' answers at measurement time one and two. Check for the very unlikely event of absolutely identical answers.

It may be practically difficult, in the harsh reality of some research settings, to fulfill the requirements of this long list. However, it is believed that quality control of the data collection process is crucial to ensure research credibility. Researchers should not only stress the ethical obligations of our interviewers; but also keep our own ethical commitment in mind. Without making sure that the data derived is accurate, readers, clients, students, and most of all, researchers are betrayed. A painful lesson had been learned, which has been used constructively for this article and for the design of our present research in Zimbabwe. There are certainly many future research opportunities in the field of interview quality assurance in general (e.g., investigation of the full-time interviewer personality as called for by Durant, 1946) and in particular (e.g., in our case the impact of Shona culture in interviewer behavior and the morality effects of the current Zimbabwean socio-political context) that are waiting to be discovered. However, it is hoped that this literature review and these experiences will engender a degree of caution in future research and, ultimately, lead to an increase in the trustworthiness of interview-based data.

References

- Biemer, P.P. & Stokes, L.S. (1989). The optimal design of quality control samples to detect interviewer cheating. *Journal of Official Statistics*, 5(1), 23–39.
- Bennett, A.S. (1948). Survey on problems of interviewer cheating: Observations of the so-called cheater problem Among Field Interviewers. *International Journal of Opinion and Attitude Research*, 2, 89–96.
- Case, P. B. (1971). How to catch interviewer errors. *Journal of Advertising Research*, 11(2), 39–43.
- Crespi, L. P. (1946). Further observations on the cheater problem. *The Public Opinion Quarterly*, 10, 646–649.
- Crespi, L. P. (1945). The cheater problem in polling. *The Public Opinion Quarterly*, 9, 431–445.
- Durant, H. (1946). The cheater problem. *The Public Opinion Quarterly*, 10, 288–291.
- Elinson, J., & Cisin, I. H. (1948). Detection of interviewer cheating through scale techniques. *The Public Opinion Quarterly*, 12, 325.
- Evans, B. (1961). On interviewer cheating. *The Public Opinion Quarterly*, 25, 126–127.

- Harrison, T. (1947). A British view on cheating. *The Public Opinion Quarterly*, 11, 172–173.
- Kiecker, P., & Nelson, J.E. (1996). Do interviewers follow telephone survey instructions? *Journal of the Market Research Society*, 38(2), 161–176.
- Kinnear, T.C., & Taylor, J.R. (1979). *Marketing research – An applied approach*. Tokyo: McGraw-Hill Kogakusha Ltd.
- Roth, J. (1966). Hired-hand research. *American Sociologist*, 1, August, 190–196.
- Stewart, N., & Flowerman, S. H. (1951). An investigation of two different methods of evaluation of interviewer job performance. *Personnel Psychology*, 4, 161–170.

