Lying Takes Time: Predicting Deception in Biodata Using Response Latency

Michael A. McDaniel

Booz•Allen & Hamilton Inc.

Howard Timm

Pebble Beach, California

The opinions expressed in this paper are those of the authors and do not necessarily reflect those of their employers. Feedback on this paper would be appreciated. Correspondence may be addressed to the senior author at Booz•Allen & Hamilton Inc., 4330 East West Highway, Bethesda, MD 20814.

ABSTRACT

RESPONSE LATENCY AS AN INDICATOR OF DECEPTION

This study evaluated the hypothesis that it takes longer to lie than to respond truthfully to biodata items. Subjects completed a 131 item computer-administered biodata instrument under operational and "ground truth" instruction sets. For all items, lying subjects took longer to respond to an item than did truthful subjects. The difference in response latencies was moderated by item content. The largest differences were found for illegal drug use and employment history items. The smallest difference was found for social desirability items. Although the study has several limitations, this research suggests that response latency is a promising method for detecting response distortion in biodata.

The study of reaction time or response latencies was one of the first topics of research in psychology. Beginning in the 1860s, the measurement of reaction time was a central research topic in the laboratories of both Sir Francis Galton and Wilhelm Wundt. As noted by Jensen (1985)[1], these seminal psychological researchers pursued reaction time research for quite different purposes. Galton's primary interest was in the nature and measurement of individual differences. In Galton's laboratory, reaction time was one of many measures of individual differences in human faculties. Galton hoped that some weighted combination of a variety of measures would provide an objective measure of general mental capacity and provide a prediction of future success in life. In contrast, Wundt viewed individual differences in reaction time as error variance. Wundt viewed reaction time as the primary technique for the decomposition of mental activity including such processes as perception, cognition, and judgment.

In more recent times, reaction time measurement has been an important research tool in experimental cognitive psychology. Here, the emphasis is on using response latency in understanding cognition and not an assessment of sensory-motor ability. Jensen (1985) notes that reaction time can be decomposed into sense organ lag, peripheral nerve transmission time, muscle latency, and brain time. Jensen summarizes the literature as indicating that cerebral activity and not sensory motor mechanisms is the main source of delay in response latency. Jensen notes that a general assumption in cognitive-oriented response latency research "is that information

---

[1] Jensen, A. R. (1985). Methodological and statistical techniques for the chronometric study of mental abilities. In C. R. Reynolds and V. L. Wilson (Eds.). Methodological and statistical advances in the study of individual differences. New York: Plenum.

processing takes place in real time in a sequence of stages and that the total measured time from the initiation of a mental task can be analyzed in terms of the time required for each stage" (p. 56).

The present research seeks to apply this assumption of response latency research to a pressing problem in personnel selection: response distortion in personnel selection. Many personnel selection instruments are designed to assess deviant or undesirable behaviors of a person's life that may impact on the person's future job performance. Such measures include the paper-and-pencil honesty tests, Hogan's personnel reliability measure, Inwald's Personnel Selection Inventory and Gough's Personnel Reaction Blank. Less psychometrically elegant measures such as application blanks and employment interviews also request information about a person's past employment and life history experiences. When the questions are very obtrusive in their intent (Have you ever stolen anything from your employer?), legitimate questions arise concerning the intentional distortion of responses by applicant's who seek to hide past undesirable behaviors from a potential new employer.

The hypothesis of the present study is that it takes longer to lie to an item concerning past behavior then it does to tell the truth. Consider the presentation of an item requiring a true or false response: "I have gotten into serious trouble because of cocaine use." For one who truthfully answers "false" to the question, the person's reaction time could be decomposed into stages relating to comprehending the stimulus, identifying the correct response, evaluating the correct response against a standard of an "acceptable" response, and preparing and executing a response of "false." However, a person who has experienced serious trouble due to cocaine use and who recognizes

that an admission of such behavior will not constitute a desired response, must spend some additional time formulating the decision to lie.

Method

Instrument.

An 131 item survey was administered. All items required a true or false response. The items included the 68 item Hogan Reliability Scale[2], a 13 item social desirability scale (e.g., unlikely virtues) from an unpublished instrument, and 22 items from the Inwald Selection Inventory[3] that focussed on drugs and criminal activities.

Subjects.

Subjects were 224 students taking classes in criminal justice. The subjects were volunteers who participated in the project to earn extra credit. Many of the students aspired to work as police officers.

Instruction Manipulation.

The instrument was administered twice in the same testing session. When the subject was first presented with the survey, the computer presented instructions stating that the instrument was being evaluated as a tool for screening applicants for security related positions. The instructions were:

> Please complete the following personnel security survey. This survey is an initial draft of a screening instrument that may be used to select applicants for sensitive positions in both government and industry. After completing the survey you will be given your actual score and informed how that

---

[2] Hogan, J. & Hogan, R. (1989). How to measure employee reliability. Journal of Applied Psychology, 74, 273-279.

[3] Inwald, R. Manual for the Inwald Selection Inventory. Kew Gardens, NY: Author.

> score compares with others who have answered the same survey when they were competing for a position. THEREFORE, PLEASE ANSWER THE QUESTIONS EXACTLY THE SAME WAY YOU WOULD IF YOU WERE TAKING THE TEST FOR A JOB THAT YOU WERE INTERESTED IN OBTAINING. Otherwise the feedback that you receive will not be meaningful to you.

Subjects were then presented with instructions concerning responding to all questions with either "true" or "false." After completing the survey, the subjects were provided feedback on scales derived from the Inwald items based on norms derived from earlier research projects. The subjects were then presented with the "ground truth" instructions:

> Now that you have completed the survey and received your feedback, please go through the same items again. This time, however, answer the questions as honestly as possible. You will not be asked for your name or any other information that could ever be used to link you to your responses. This information will be used solely to establish ground truth and to determine how common certain types of behavior really are throughout our society today.

After completing the survey under the ground truth condition, the subjects responded to several demographic items.

Analyses.

Random Response Screen. A random response scale was developed to identify subjects who were not responding with care to the survey. The scale was incremented by a value of one for aberrant score responses based on several variables (number of responses under one second, number of 2-alike though 6-alike responses in a sequence, number of true to false and false to true changes, number of true responses, number of false responses). For each variable contributing to the random response scale, an

aberrant response was defined as a response lying in the top and/or bottom two percent of the response distribution. Of the 224 observations, 26 were dropped based on the random response screen, leaving 198 cases for the remainder of the analyses.

Score Standardization. Response latency scores were expressed in three scales, raw latencies and two standardization approaches designed to control for individual differences in response time. Raw score latencies were calculated by recording the time in seconds between the presentation of the item on the computer screen and the response. The first standardization approach, expressed the response latency in deviations from the subjects own median response time across items. The median deviation was judged more useful than mean deviations because the mean is more subject to distortions due to extreme response latencies (e.g., a long latency could be caused by the subject taking a short break before answering the item). The second standardization approach standardized for both the subjects own median response time and the median response time for each item across all subjects. For the sake of brevity, only the results for the raw score latencies are presented. The conclusions are identical regardless of whether the raw scores or one of the standardized scores are employed.

Results

Table 1 presents the results of a general linear model analysis. For this analysis, the unit of observation is the item. Thus, there are 131 observations. The dependent variable is the median reaction time for the item. The independent variables are:

- The percent of people who lied on the item.
- The content category of the item divided into 7 categories:

- Drug use items (12 items)
- Employment history items (12 items)
- Exhibitionism items (12 items; e.g., "I like to be the center of attention")
- Alcohol use items (12 items)
- Legal/Crime items (10 items)
- Social desirability items (13 items)
- All other items (65 items).

The model shown in Table 1 includes two main effects (Percent who lied, Item content) and their interaction. Table 1 indicates that the full model is statistically insignificant but that the main effect for lying is highly significant. The interpretation is as follows: The greater the proportion of people who lied on the item, the greater the response time.

The analysis presented in Table 1 is not ideal for two reasons. First, it does not directly test the hypothesis that it takes longer to lie than to respond truthfully. Second, the unit of analysis is an item. This restricts the "sample size" to 131 items and results in a low power test.

Table 2 compares the response times of those who responded honestly to the items with those who responded dishonestly. Each row of the table provides the results for sets of items (for example, all items, drug use items, etc.). For each item, the median response time was calculated separately for two groups of people, the honest and dishonest responders. The response latency presented in the column labelled "Honest" for the first row of data ("All items") is the mean of the 131 response latency

medians for the honest responders. The response latency presented in the column labelled "Dishonest" for the first row of data ("All items") is the mean of the 131 response latency medians for the dishonest responders. The value in the column labelled "Difference" is the honest mean time subtracted from the dishonest mean time. The value in the column labelled "d" is the difference between the mean response latency medians of honest and dishonest responders expressed in standard deviation units (Cohen's *d*).

Table 2 shows that the average response time across all items for honest responders was 4.03 seconds. Dishonest responders, on average, took 4.63 seconds to answer the item. Thus, on average, dishonest responders took .60 seconds longer to respond than did honest respondents. Since the interaction effect of the general linear model analysis summarized in Table 1 was not significant, that analysis failed to support the inference that item content moderated the magnitude of the response latency differences between the honest and dishonest responders. However, Table 2 suggests that the magnitude of the response latency differences meaningfully covaries with item content. Those who lied on the drug use items took 1.10 seconds more time to respond than did the honest responders. Likewise, those who lied on employment items took .94 seconds longer to respond. In contrast those who lied on social desirability items required only .23 seconds longer to respond than did the honest responders.

Discussion

This study strongly supported the hypothesis that it takes longer to lie in response to biodata items than it does to tell the truth. For this set of 131 true-false

items, honest responders took 4.03 seconds, on average, to respond while dishonest responders required an additional .60 seconds. Although a response time difference of about one-half second may appear to some to be a small effect, it is actually substantial compared with the 4 second average. For the example of the set of all items, if the honest responders are at the 50th percentile of response latency, the dishonest responders are at the 62nd percentile.

For the analyses concerning differences in response time as moderated by item content, the results are mixed. The general linear model analysis suggests no moderating effect (as indicated by the nonsignificant interaction effect). However, the pattern of response latency differences in Table 2 is compelling. Based on Table 2, it appears that lying on some item content takes longer than lying on other item content. In particular, lying on drug use and employment history items appears to be the most readily detected by response latency analysis. In contrast, honest and dishonest responders have approximately the same response latency on social desirability items.

Although this study has provided useful information concerning the value of response latency for detecting response distortion, the study suffers from several limitations. First, the analyses presented here do not address the effectiveness of a response latency scoring scheme in an operational program to differentiate liars from truthful responders. Second, the definition of "ground truth" is a major source of controversy in any type of deception detection research (see, for example, the literature on polygraph research). One has no way of ascertaining that the responses obtained in the ground truth condition are reliable and valid measures of ground truth. Third, the subjects did not know that their response latencies were a topic of study. In an

operational setting, subjects with awareness that a response latency measure is used to detect deceit may develop strategies to subvert the predictive value of the latency measurement.

Although this study suffers from several limitations, the results are relatively clear in their support of the hypothesis. It takes longer to lie than to respond truthfully. We argue that further research in this area is warranted.

Table 1

General Linear Model Analysis
Response Latency as a Function
of Percent Who Lied and Item Content

Dependent Variable: Response Latency

Independent Variables: Percent Who Lied on Item (Variable name = LIE)
Item Content Category (Variable name = IC)

Source	*DF*	*SS*	*MS*	*F*	*p*	R^2
Model	15	27.02	1.80	1.41	0.15	.16
Error	115	146.45	1.27			
Corrected Total	130	173.47				

Source	*DF*	*SS*	*F*	*p*
LIE	1	10.83	8.50	0.0043
IC	7	11.94	1.34	0.2383
LIE*IC	7	4.25	0.48	0.8491

Table 2

Response Times For Honest And Dishonest Responders By Item Type.

Item Type	Honest	Dishonest	Difference	d
All Items (131 Items)	4.03	4.63	0.60	.29
Drug Use Items (12 items)	3.42	4.52	1.10	.55
Employment History Items (12 items)	4.63	5.57	0.94	.42
Exhibitionism Items (7 items)	4.21	5.01	0.80	.37
Alcohol Use Items (12 items)	4.07	4.67	0.60	.28
Legal/Criminal Items (10 items)	4.15	4.61	0.46	.22
Social Desirability Items (13 items)	4.12	4.35	0.23	.10