

2021

## Identifying Faking on Forced-Choice Personality Items Using Mouse Tracking

Irina Kuzmich

*Baruch College & the Graduate Center, City University of New York*

Charles Scherbaum

*Baruch College & the Graduate Center, City University of New York*

Follow this and additional works at: <https://scholarworks.bgsu.edu/pad>

 Part of the [Human Resources Management Commons](#), [Industrial and Organizational Psychology Commons](#), and the [Other Psychology Commons](#)

[How does access to this work benefit you? Let us know!](#)

---

### Recommended Citation

Kuzmich, Irina and Scherbaum, Charles (2021) "Identifying Faking on Forced-Choice Personality Items Using Mouse Tracking," *Personnel Assessment and Decisions*: Number 7 : Iss. 1 , Article 5.

DOI: <https://doi.org/10.25035/pad.2021.01.005>

Available at: <https://scholarworks.bgsu.edu/pad/vol7/iss1/5>

This Main Article is brought to you for free and open access by the Journals at ScholarWorks@BGSU. It has been accepted for inclusion in Personnel Assessment and Decisions by an authorized editor of ScholarWorks@BGSU.

# IDENTIFYING FAKING ON FORCED-CHOICE PERSONALITY ITEMS USING MOUSE TRACKING

Irina Kuzmich<sup>1</sup> and Charles Scherbaum<sup>1</sup>

1. Baruch College & the Graduate Center, City University of New York

## ABSTRACT

### KEYWORDS

faking, mouse tracking, impression management, personality

This research utilizes mouse tracking as a potential behavioral method to examine cognitive processes underlying faking on forced-choice personality inventories. Mouse tracking is a method from social categorization research that captures a variety of metrics related to motor movements, which are linked to cognitive processing. To explore the utility of this method, we examined differences in the mouse tracking metrics of those instructed to respond honestly or to fake. Our findings show that there is a distinguishable difference in the behavioral response of those who are faking when responding to pairs of personality descriptors presented in a forced-choice format compared to those who are responding honestly. Implications and contributions of this study include insights into the cognitive processing that can occur while responding to personality items when respondents are faking and a demonstration of how mouse tracking methods can be used to detect faking.

Despite the benefits of personality assessment in the workplace, the major criticism is that individuals can fake their responses to personality items (Morgeson et al., 2007). The term faking describes attempts to respond to personality items in a way that portrays the individual more favorably and increases the chances of receiving desired outcomes (Ziegler et al., 2012). Given the potential implications of inaccurate personality assessment in selection contexts, researchers have proposed a number of theoretical models to understand who will fake, when they will do it, and how they do it.

Many of these theoretical models have focused on the decision-making and cognitive processes underlying faking (e.g., Ellingson & McFarland, 2011; Goffin & Boyd, 2009; Griffith et al., 2011; Holden et al., 1992; Komar et al., 2010; König et al., 2012; Kuncel et al., 2011; Robie et al., 2007; Shoss & Strube, 2011; Vasilopoulos et al., 2000). These models attempt to peer inside what Ziegler (2011) has called the “black box” of faking to understand the cognitive processes operating when selecting a response option and if there are different cognitive processes operating when individuals are responding honestly or faking.

Research examining cognitive processes underlying faking has generally relied on two research methodologies. First, studies have employed qualitative methods in which individuals are asked to explain their cognitive processes while responding to the personality items or asked to

explain their response choices (e.g., König et al., 2012; Kuncel & Tellegen 2009; Robie et al., 2007). This research has generally found that individuals report different cognitive processes when faking compared to when responding honestly. Second, studies have used behavioral methods that are theoretically associated with cognitive processes including eye tracking (e.g., van Hooft & Born, 2012) and item response time (e.g., Fine & Pirak, 2016; Holden et al., 1992; Komar et al., 2010; Shoss & Strube, 2011). This research has generally found that different behavioral patterns that are theoretically associated with the cognitive processes are displayed by those faking compared to those responding honestly. However, the research is less consistent on the nature of some of these differences. For example, the research on item response time has produced conflicting findings showing that those who fake take longer to respond in some studies and take less time to respond in other studies (Dilchert & Ones, 2012). Despite the intuitive appeal of behavioral methods for examining faking, their viability is an open question based on the current state of the research.

The purpose of this study is to examine a novel behavioral method for examining cognitive processes underlying

---

Corresponding author:  
Charles Scherbaum  
Email: Charles.Scherbaum@baruch.cuny.edu

faking called mouse tracking. Originally developed in the social categorization literature, mouse tracking can be used to study cognitive processes by tracking the motor movement of the hand using a computer mouse as an individual chooses among different response options (Freeman & Ambady, 2010). As described in the subsequent sections, mouse tracking produces several behavioral metrics beyond response time that may be useful for examining faking. Additionally, this paper examines the usefulness of mouse tracking with forced-choice personality items. The majority of the previous research using behavioral methods to study cognitive processes underlying faking has used single statement personality items with rating scales. Thus, our knowledge of the usefulness of behavioral methods with forced-choice personality items is limited. Given the increasing popularity of forced-choice personality inventories (Cao & Drasgow, 2019; Zhang et al., 2020), research focusing on applying behavioral methods for examining cognitive processes underlying faking to forced-choice personality items is warranted.

To examine the utility of mouse tracking with forced-choice personality items, this study uses a between-subjects experimental design with an instructed response manipulation. In the following sections, the previous research on cognitive processes underlying faking is reviewed, the mouse tracking methodology is described, and the research hypotheses are presented.

### Prior Research on Cognitive Processes Underlying Faking

Researchers have long been interested in the decision-making and cognitive processes that individuals engage in when faking personality items. A variety of theoretical models have been offered that describe the different decision-making and cognitive processes operating when individuals respond honestly or fake (e.g., Ellingson & McFarland, 2011; Goffin & Boyd, 2009; Griffith et al., 2011; Holden et al., 1992; König et al., 2012; Komar et al., 2010; Kuncel et al., 2011; Robie et al., 2007; Shoss & Strube, 2011; Vasilopoulos et al., 2000). A number of these models have focused on three different cognitive comparison processes between the meaning of a response to an item and a schema of the self (e.g., Shoss & Strube, 2011; Vasilopoulos et al., 2000).

One of these models has been called the self-schema model, which states that individuals have a schema of their true selves and decide to respond in a way that is contrary to that schema. A second variation of these models has been called the adopted-schema model, which states that individuals access the schema of an ideal individual and respond according to that schema. A third variation has been called the semantic-exercise model, which states that individuals evaluate each item in terms of which would be the most socially desirable response and not in reference to the self. Although these models offer different perspectives on the cognitive processes that occur when individuals are

responding honestly or faking, they share a similarity in relying on behavioral methods to capture these cognitive processes. Research has employed both eye tracking (e.g., van Hooft & Born, 2012) and item response time (e.g., Fine & Pirak, 2016; Holden et al., 1992; Komar et al., 2010; Shoss & Strube, 2011). However, the majority of the research has focused on item response time (Dilchert & Ones, 2012). This research has also focused on single statement personality items with rating scales. To our knowledge, there has been no research using behavioral methods with forced-choice personality items despite their increasing use (Cao & Drasgow, 2019; Zhang et al., 2020).

Each of these cognitive process models of faking proposes conflicting hypotheses about whether item response time should be faster or slower when individuals are faking. The collective findings of the research testing these cognitive process models are generally inconclusive (Fine & Pirak, 2016). Dilchert and Ones (2012) outlined three different conclusions that have been reached by various researchers regarding item response times: those who fake take longer to respond, those who fake take less time to respond, and the amount of time depends on the consistency between the meaning implied by the response and the schema the individual is operating from. More recently, eye tracking has been introduced as a method that may provide additional insights into the cognitive processes underlying faking. To date, only one study has used eye tracking (van Hooft & Born, 2012). Despite its promise, the lack of use likely reflects the challenges of using eye tracking and the availability of the specialized equipment (Meißner & Oll, 2018).

The purpose of this study is to examine a novel behavioral method for examining cognitive processes underlying faking called mouse tracking. Mouse tracking is an approach to studying cognitive processes using motor movements of the hand that provides a variety of behavioral metrics beyond reaction time (Freeman & Ambady, 2010). Thus, it could be well suited to studying the cognitive processes involved in faking.

### Mouse Tracking as a Behavioral Method for Studying Faking on Forced-Choice Personality Items

Mouse tracking was initially developed in the social categorization literature to capture an individual's cognitive processing during categorization tasks through hand movements while using a computer mouse (Freeman, 2009). It measures categorization processes by allowing a stimulus to be presented and then recording the response time and response trajectories, or paths, of the mouse movements from a starting point at the bottom of the screen to one of the two response options that represent different categories at the top of the screen. One response option is located in the top left corner of the screen, and the other is located in the top right corner of the screen (see Figure 1).

Consistent with continuous models of categorization, competing categories are activated, or triggered, when an

individual is presented with a stimulus, which means that multiple categories are initially partially activated. Freeman et al. (2008) demonstrated support for this partial activation of multiple categories with studies examining the categorization of sex-atypical individuals as “male” or “female” using mouse tracking. In their studies, individuals were presented with male and female faces that differed in whether they had sex-typical or sex-atypical features. For example, one study manipulated hair length as a way to examine the categorization of sex-typical and sex-atypical faces. The sex-atypical faces were those where men had long hair or women had short hair. For each trial, when a face was presented, the word “male” was in one corner of the screen and word “female” in the other.

Participants were asked to categorize each face by moving the mouse from the starting point to the correct label. An ideal trajectory of the mouse would be a straight line from the starting point to the selected response option. This representation of the ideal trajectory has been theoretically determined as the result of only one category being activated whereas the other is not. Based on the concept of motor movements as reflections of cognitive processing, it has been established by researchers utilizing mouse tracking that if only one category were to be activated, rather than multiple categories, movement from an initiation point to the category would mimic a straight line connecting these two points. This is believed to be a universal standard

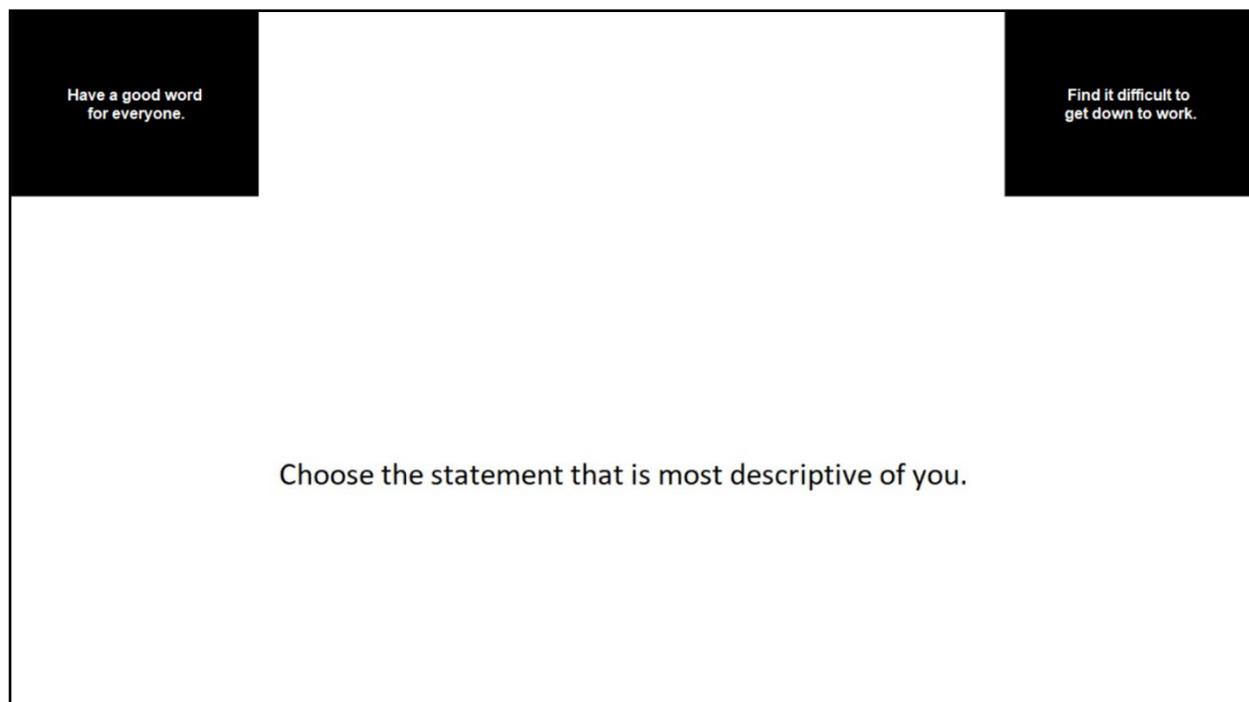
for all individuals who are engaged in a categorization task. For faces with sex-typical hair, the mouse trajectories from the starting point to the correct label more closely approximated but were not identical to an ideal trajectory. Some curvature is still present, indicating a slight partial activation of the opposite category. In contrast, as was expected, for faces with sex-atypical hair, there was greater activation of and attraction toward the opposite sex category than for faces with sex-typical hair. This was demonstrated in the form of more curved mouse trajectories that showed movement toward the opposite category before selection of the correct category. This means that these trajectories showed greater movement away from the ideal trajectory than did the trajectories for sex-typical faces.

When applied to the three common cognitive process models of faking, mouse tracking may allow for additional insights beyond item response time. For example, one would expect to see trajectories that deviate from a straight line if individuals are responding in a way that is contrary to their self-schema (self-schema model). For items that are high in social desirability, one would expect to see trajectories with little deviation from a straight line if the processes articulated in the adopted-schema model are operating. In both cases, the response time may be similar, but the behavioral response measured by the mouse movements would be different.

Most research utilizing mouse tracking has examined a

### FIGURE 1.

Example of how the screen appeared to participants for a single trial in the Runner program of the MouseTracker software package



number of metrics that can be extracted from the behavioral data of the mouse movement. The most commonly used metrics are described below.

*Area under the curve.* As previously noted, an ideal response trajectory would be a straight line from the starting point of the mouse to the selected response option if only one category, rather than multiple categories, were to be activated. Much of the work using mouse tracking examines situations where multiple categories are activated. Thus, the response trajectories are generally not a straight line. The area between actual response trajectory and the ideal response trajectory from initiation to selection of the response is used to compute the area under the curve. Larger areas under the curve indicate greater cognitive processing when selecting between two response options. For example, if two response options that are equally socially desirable are presented to an individual who is faking, one would expect greater cognitive processing to determine which option to select and a larger area under the curve than for an individual responding honestly.

*Maximum deviation.* Related to the area under the curve is the maximum value of the difference between the actual response trajectory and the ideal response trajectory. Larger maximum deviations indicate greater cognitive processing when individuals are choosing between two response options.

*X-flips.* X-flips are back-and-forth movements of the mouse between the categories along the x axis. These flips are indicative of difficulty in determining the proper response from the two options, or which of the two partially activated response options should be more fully activated and which should be deactivated. For example, Freeman et al. (2008) found that x-flips occurred more frequently when individuals are categorizing racially ambiguous faces (i.e., faces possessing a mixture of characteristics specific to White and to Black individuals) than faces that were not racially ambiguous. Due to this ambiguity, there was greater competition between the “White” and “Black” categories, and participants’ mouse movements vacillated between the two categories for the racially ambiguous faces. In the context of faking, one might expect more x-flips when individuals are choosing between response options that are equal in social desirability.

*Response time.* Similar to prior research on faking, the total response time to each item can be recorded. In mouse tracking, the response time can be decomposed into time from presentation of the response options to initiation of the mouse movement and time from initiation of movement to response selection.

*Velocity.* Velocity is a measure of the distance per second that the mouse is moved from the starting point to the response option. It is recorded as distance moved between different time-points throughout the response process. When an individual is first presented with a stimulus and two response options, the response options enter into and compete with one another within the processing pipeline.

This competition leads to a slower velocity of the mouse movement. Once the competition is resolved and one response overtakes the other, there is a sharp increase in the velocity due to the cessation of competition. This is followed by a sharp decrease in the velocity as the response option is approached. For example, Duran et al. (2010) found differences in the velocity of response patterns for those who were telling the truth and those who were lying such that when answering truthfully, the point of the initial increase in velocity and the subsequent point of decrease in the velocity as the responses were approached occurred earlier than when false answers were given.

### Hypotheses

Hypotheses can be offered for the response trajectories of those who are faking and those responding honestly. Monaro et al. (2017) conducted a study in which half of the participants responded to biographical questions according to their true identity, whereas the other half answered based on an assumed identity. Using mouse tracking, they were able to compare the differences in response trajectories of “truth tellers” and “liars.” Liars, while answering according to a false identity, had a greater maximum deviation from the ideal trajectory than did truth tellers, as well as a greater area under the curve. Consistent with this research, we expect that those who are faking will have larger area under the curve and larger maximum deviation than those who are responding honestly.

**Hypothesis 1:** Individuals who are faking will have a larger area under the curve compared to individuals responding honestly when responding to the forced-choice personality items.

**Hypothesis 2:** Individuals who are faking will have a larger maximum deviation compared to individuals responding honestly when responding to the forced-choice personality items.

When presented with two response options that do not differ in social desirability, those engaging in faking will not be able to easily categorize one response as more socially desirable than the other. Therefore, they will demonstrate indecision as to which response option to select. This will be seen as repeated movements toward and then away from each of the response options in the form of x-flips. Individuals who are responding honestly may still demonstrate some indecision with such item pairs, but it will be less frequent than for those who are faking as they are not trying to ascertain the more socially desirable response.

**Hypothesis 3:** For forced-choice item pairs that are equal in social desirability, those who are faking will exhibit a greater number of x-flips in their response trajectories than those who are responding honestly.

Consistent with the work of Duran et al. (2010), we expect differences in response time and velocity for those who are faking compared to those responding honestly. Specifically, we expect that those who are faking will take longer to respond, spend more time at the initiation point, take longer when moving from the initiation point to the response choice, and have slower velocities that peak at later points than those who are responding honestly.

**Hypothesis 4a:** Those who are faking will have total response times that are longer than will those who are responding honestly to the forced-choice personality items.

**Hypothesis 4b:** Those who are faking will spend more time at the initiation point than will those who are responding honestly to the forced-choice personality items.

**Hypothesis 4c:** Those who are faking will spend longer moving from the initiation point to the response choice than will those who are responding honestly to the forced-choice personality items.

**Hypothesis 5a:** Those who are faking will have initial increases in velocity that occur later than those who are responding honestly to the forced-choice personality items.

**Hypothesis 5b:** Those who are faking will have decreases in velocity occur later than those who are responding honestly to the forced-choice personality items.

## METHOD

### Participants

The participants in this study were 136 undergraduate students recruited from a Northeastern U.S. university. About half of the participants were female, and close to half identified as Asian (see Table 1 for participant demographics). Only right-handed participants were used in this study as handedness can have an impact on mouse movements and trajectories (Hehman et al., 2015; Spivey et al., 2005).

### Design

This study used an experimental design in which response instructions were manipulated. Participants were randomly assigned to a faking response condition or an honest response condition before completing the forced-choice personality items. Manipulating faking through response instructions has been effectively used in many studies on faking behavior, and response instructions have been shown to induce faking (Mesmer-Magnus & Viswesvaran, 2006).

The instructions used in this study are a slightly-modified version of those used by Scherbaum et al. (2013). Those in the honest condition received the following instructions:

Please complete this personality inventory as honestly as you can. The results will be completely anonymous and will be used for research purposes only. It is very important that you respond to this survey by describing yourself as you really are, not as you wish to be.

Those in the faking condition received the following instructions:

Please complete the personality inventory as if you were applying for a job you really want. To increase your chances of being hired, you should respond in ways that will make you look like the ideal job candidate.

### Measures

*Personality.* A forced-choice personality inventory was developed using single statement items from the International Personality Item Pool (IPIP). To determine the perceived desirability of each IPIP item, we conducted a pilot study through Amazon Mechanical Turk in which participants rated the IPIP items on their level of social desirability. The pilot study included 68 participants who were paid \$0.50 for completing the study. All of the participants were located in the United States and had HIT approval ratings that were greater than 95%. Participants rated 100 single statement personality items from the IPIP that were selected to mimic the NEO-PI-R with 20 items for each of the five personality factors. Participants rated these 100 items on the likelihood it would be endorsed by someone who wants to make the best impression when applying for a job using anchors ranging from extremely unlikely (0) to extremely likely (4). Based on the ratings, items were classified as “low,” “medium,” or “high” in desirability. Items low in desirability were those with average ratings less than or equal to 0.75. Items classified as medium in desirability were those whose average ratings ranged from 1.42 to 2.68. Items high in desirability were those whose average ratings were greater than or equal to 3.50. A wider range of desirability ratings had to be used to classify items as medium in social desirability as there were few items that had average ratings close to the exact middle of the scale (a rating of 2). Consequently, a narrower range was used for classifying items as low in desirability or high in desirability. The cut offs for the low and high desirability items were chosen in such a way that 20 of the single statement items would be classified as high in desirability and 20 of the single statement items would be classified as low in desirability. The remaining 60 single statement items would be classified as medium in desirability.

**TABLE 1.**  
Frequencies and Percentages of Demographic Characteristics of Participants

| Characteristic                | <i>N</i> | %    |
|-------------------------------|----------|------|
| Gender                        |          |      |
| Male                          | 64       | 47.1 |
| Female                        | 72       | 52.9 |
| Race/Ethnicity                |          |      |
| Asian                         | 63       | 46.3 |
| Black or African American     | 12       | 8.8  |
| Hispanic/Latino               | 21       | 15.4 |
| White                         | 30       | 22.1 |
| Other                         | 3        | 2.2  |
| Multiple races or ethnicities | 7        | 5.1  |

Using the desirability classifications of the 100 single statement items, we created 90 forced-choice item pairs. Sixty of the 90 forced-choice item pairs were unequal in desirability (e.g., a high desirability item paired with a low desirability item). Thirty of the 90 forced-choice item pairs were equal in desirability with ten forced choice item pairs at each of the low, medium, and high desirability levels. Due to the wider range of the desirability ratings used to classify items as medium, pairings within the medium classification were made between single statement items that were closest in the desirability ratings.

*Demographic characteristics.* Participants were asked to provide information about their gender and race.

### Procedure

Participants were randomly assigned to the faking ( $n = 68$ ) or honest condition ( $n = 68$ ) at the start of the study and responded to the 90 forced-choice personality item pairs. Each forced choice item pair represented a separate trial. Participants responded to the forced-choice item pairs using the *MouseTracker* program (Freeman, 2009; Freeman & Ambady, 2010). For each forced-choice item pair, participants saw the instructions, “choose the statement that is most descriptive of you” in the center of the screen. The two response options were the paired personality statements. The statements were placed in the upper left corner and upper right corner of the screen. The cursor was at a fixed starting point, toward the bottom of the screen, at the beginning of each trial. Participants moved the mouse to select one of the two response options. They were informed that they should start moving the mouse as soon as the item pair is presented. Participants completed each trial by selecting one of the statements in the pairing (see Figure 1 for an example of how the screen appeared for a single trial).

## RESULTS

### Outliers and Data Screening

Prior to running any analyses, outlier trajectories were removed from the data. Each trajectory reflects one trial.

Outliers were trajectories for which the value on the area under the curve metric fell more than three standard deviations above or below the mean. This resulted in the removal of 314 trajectories, or 2.6% of the total 12,240 trajectories. Additionally, approximately 1.0% of the total trajectories displayed aberrant movement (e.g., zig-zags) and were removed due to being uninterpretable (see Figure 2 for examples of such trajectories). Descriptive statistics and correlations between the metrics are reported in Tables 2 and 3.

### Tests of Hypotheses

Hypothesis 1 proposed that those in the faking condition would have trajectories with a larger area under the curve than those in the honest condition. Those in the faking condition demonstrated a statistically significant larger area under the curve ( $M = 2.10$ ,  $SD = 1.13$ ) than those in the honest condition ( $M = 1.74$ ,  $SD = .930$ ),  $t(134) = 2.09$ ,  $p = .038$ ,  $d = .356$ . Thus, Hypothesis 1 was supported (see Figure 3 for representations of the average trajectory by condition).<sup>1</sup>

Hypothesis 2 proposed that those in the faking condition would have larger maximum deviations than those in the honest condition. Those in the faking condition demonstrated a slightly larger maximum deviation ( $M = .773$ ,  $SD = .316$ ) than those in the honest condition ( $M = .705$ ,  $SD = .267$ ), but the difference was not statistically significant,  $t(134) = 1.39$ ,  $p = .168$ ,  $d = 0.236$ . Thus, Hypothesis 2 was not supported.

Hypothesis 3 proposed that those in the faking condition would have a greater number of x-flips than those in the honest condition on forced-choice item pairs that were equal in desirability. Those in the faking condition exhibited a slightly larger but statistically nonsignificant number of x-flips ( $M = 7.54$ ,  $SD = 1.63$ ) compared to those in the honest condition ( $M = 7.34$ ,  $SD = 1.60$ ) on items that were equal in desirability,  $t(134) = .714$ ,  $p = .476$ ,  $d = .124$ . Thus, Hypothesis 3 was not supported.

Hypothesis 4 offered predictions about differences in the time to respond between the two conditions. Hypothesis 4a predicted that total response time would be greater for those in the faking condition than for those in the honest condition. The difference between those in the faking ( $M = 2982.38\text{ms}$ ,  $SD = 1005.43\text{ms}$ ) and honest condition ( $M = 2870.19\text{ms}$ ,  $SD = 1127.65\text{ms}$ ) were in the correct direction; however, the difference was not statistically significant,  $t(134) = .612$ ,  $p = .541$ ,  $d = .11$ . Hypothesis 4b stated that those in the faking condition would spend more time at the initiation point than would those in the honest condition. The difference between those in the faking ( $M = 213.40\text{ms}$ ,  $SD = 105.07\text{ms}$ ) and the honest condition ( $M = 208.48\text{ms}$ ,

<sup>1</sup> For ease of representation, trajectories in Figure 3 were re-mapped rightward. The ideal trajectory is included for comparison of the two response curves.

**TABLE 2.**  
Means and Standard Deviations for Key Metrics by Condition

| Metrics                  | Condition                  |                            |
|--------------------------|----------------------------|----------------------------|
|                          | Faking<br>( <i>n</i> = 68) | Honest<br>( <i>n</i> = 68) |
| Maximum deviation        | .773 (.316)                | .705 (.267)                |
| Area under the curve     | 2.10 (1.13)                | 1.74 (.930)                |
| Initiation time (ms)     | 213.40 (105.07)            | 208.48 (77.99)             |
| Movement time (ms)       | 2649.57 (1165.75)          | 2768.99 (993.60)           |
| Total response time (ms) | 2982.38 (1005.43)          | 2870.19 (1127.65)          |

Note. Standard deviations appear in parentheses.

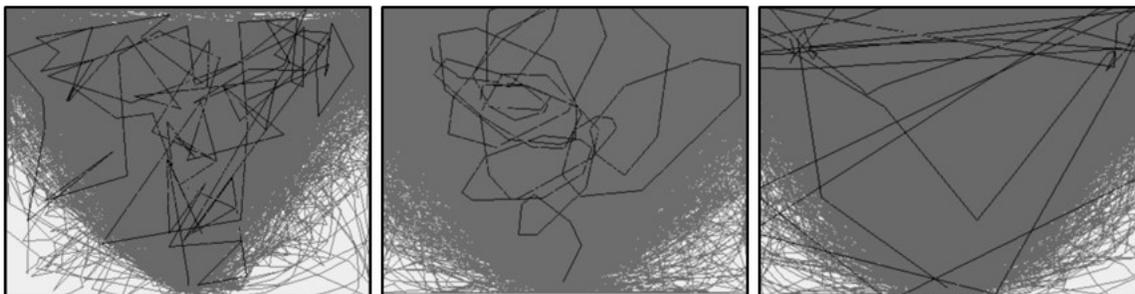
**TABLE 3.**  
Correlations Between Condition and Mouse Tracking Metrics for All Item Pairs

| Variable                | 1     | 2       | 3      | 4      | 5      | 6      | 7    |
|-------------------------|-------|---------|--------|--------|--------|--------|------|
| 1. Condition            | 1.00  |         |        |        |        |        |      |
| 2. Initiation time      | 0.03  | 1.00    |        |        |        |        |      |
| 3. Movement time        | 0.06  | -0.10   | 1.00   |        |        |        |      |
| 4. Total response time  | 0.05  | -0.01   | 0.99** | 1.00   |        |        |      |
| 5. Maximum deviation    | 0.12  | -0.20*  | 0.30** | 0.28** | 1.00   |        |      |
| 6. Area under the curve | 0.18* | -0.17*  | 0.24** | 0.23** | 0.97** | 1.00   |      |
| 7. X-flips              | 0.09  | -0.31** | 0.07   | 0.05   | 0.45** | 0.45** | 1.00 |

Note. \* $p < .05$ ; \*\* $p < .01$ . Condition: Faking = 1 and Honest = 0.  $n = 68$  in both conditions.

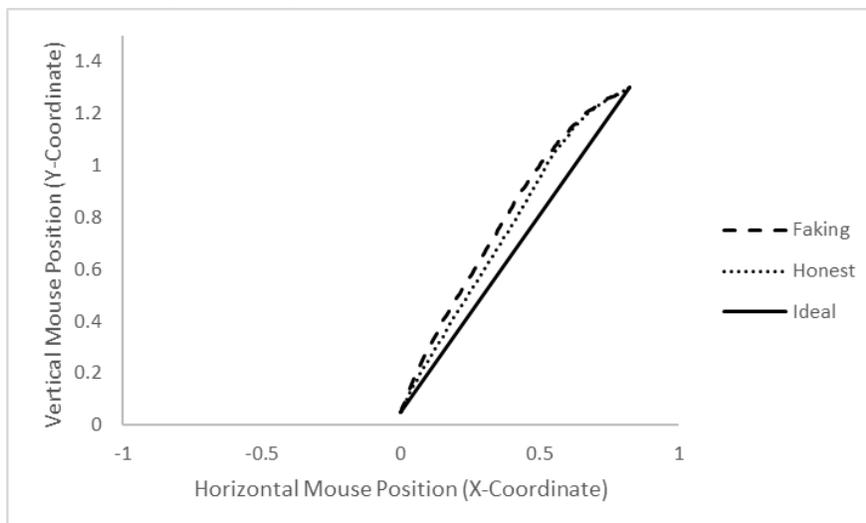
**FIGURE 2.**

Screenshots of three trajectories that exhibited looping and were removed from the data



**FIGURE 3.**

Mean mouse trajectories by condition



$SD = 77.99\text{ms}$ ) was in the correct direction; however, the difference was not statistically significant,  $t(134) = .310$ ,  $p = .757$ ,  $d = .053$ . Hypothesis 4c predicted that those in the faking condition would spend more time moving from the initiation point to the response choice than would those in the honest condition. The difference between the faking ( $M = 2768.99\text{ms}$ ,  $SD = 993.60\text{ms}$ ) and the honest ( $M = 2649.57\text{ms}$ ,  $SD = 1165.75\text{ms}$ ) conditions was in the hypothesized direction; however, the difference was not statistically significant,  $t(134) = .643$ ,  $p = .521$ ,  $d = .11$ . Looking across the three parts of Hypothesis 4, it can be seen that the standard deviations for the metrics are relatively large. No participant could be clearly identified as an outlier. Thus, Hypothesis 4 was not supported.

Hypothesis 5a and 5b proposed that there would be a difference in the timepoints at which participants in the faking condition exhibited an increase in their velocity, followed by a decrease compared to the honest condition. To examine whether this difference exists, the velocity of mouse movement across the course of choosing a response was divided into 200ms time bins to be consistent with prior research using mouse tracking (e.g., Duran et al., 2010).<sup>2</sup> Testing these hypotheses requires visual inspection of the velocity curves, which were plotted for each condition across the time bins. Figure 4 provides a plot of the velocity curves up to 3200ms as the critical changes in velocity occur in the earlier time bins. As can be seen, the increase in velocity occurs in both conditions during the 401ms–600ms time bin. There is a slightly larger increase in velocity for the faking condition for the early time bins. After the initial increase, the subsequent decrease in velocity also occurs in the same time bin (601ms–800ms) in both conditions. Thus, taken together, this pattern of changes in velocity does not provide support for Hypothesis 5a or Hypothesis 5b.

### Supplemental Analyses

The primary hypotheses were tested using item pairs that were either equal and unequal in social desirability. As a follow up test, we examined the relationship between response condition and each of the mouse tracking metrics using only the item pairs that were equal in social desirability. We expect that the metrics of mouse tracking are most useful when the forced-choice item pairs are equal in desirability. These are the situations where both response options will be partially activated and deviations from the ideal response trajectory can be observed. When considering only forced-choice item pairs that were equal in desirability, area under the curve is still correlated with condition and the magnitude of the correlation is larger (see Table 4). The

<sup>2</sup> Time bins of smaller sizes, as small as 20ms, were also created. A similar pattern emerged regardless of the chosen size of the time bin. For this reason, a size of 200ms was used for optimal graphical representation.

correlations for the other metrics showed little to no change.

One potential concern with the area under the curve metric is that it can be viewed simply as an alternative indicator of response time. To determine if area under the curve has utility beyond response time, we conducted a logistic regression analysis predicting condition from response time and area under the curve for item pairs that were equal in desirability. The overall model was statistically significant ( $\chi^2(2) = 6.103$ ,  $p = 0.047$ ). Response time was not a statistically significant predictor ( $B = 0.00$ ,  $p = 0.810$ ) of condition, but area under the curve was a significant predictor ( $B = 0.372$ ,  $p = 0.023$ ). At least in this research, area under the curve appears to offer additional insights into faking beyond reaction time.

## DISCUSSION

In conducting this study, our objective was to take a novel approach to the examination of faking on personality inventories, utilizing mouse tracking. To this end, we conducted an experimental study that manipulated faking, allowing us to compare behavioral indicators of cognitive processing between those who responded honestly and those who faked their responses. We found that there are differences on one of the behavioral indicators between those who are faking and responding honestly.

### Theoretical Implications

This study provides a window into individuals' self-reflection when responding to personality items. The introspection differs depending on whether one is engaged in faking or responding honestly according to one of the behavioral indicators. Those individuals who were engaged in faking exhibited trajectories with greater attraction to the unselected response option, reflected by the trajectory having a greater area under the curve. This demonstrates that when the two response options are in the processing pipeline, the unselected response option is more activated and poses greater competition to the response option ultimately selected by those engaged in faking when compared to those responding honestly. These effects were stronger when both items were equal in social desirability and predicted response condition even after controlling for response time. Although testing the various cognitive process models presented in Shoss and Strube (2011) was not the goal of this research, the results are consistent with the semantic-exercise model.

### Practical Implications

There are also several practical implications for our findings. Mouse tracking has the potential to support forced-choice item development and testing. For example, the desirability of item pairs needs to be considered when forming forced-choice personality inventories. Mouse

TABLE 4.

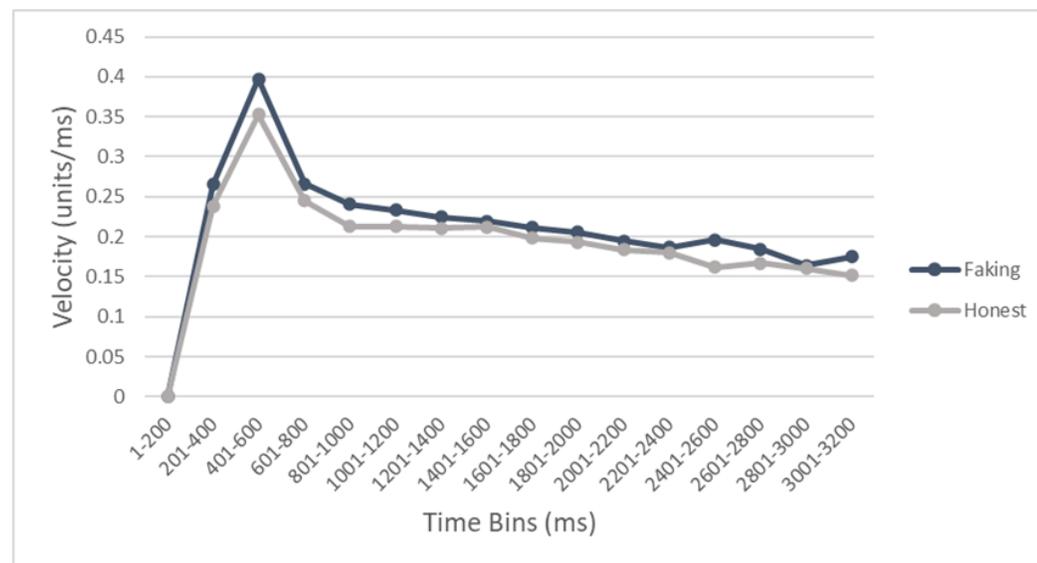
Correlations Between Condition and Mouse Tracking Metrics for Item Pairs Equal in Social Desirability

| Variable                | 1     | 2       | 3      | 4      | 5      | 6      | 7    |
|-------------------------|-------|---------|--------|--------|--------|--------|------|
| 1. Condition            | 1.00  |         |        |        |        |        |      |
| 2. Initiation time      | -0.06 | 1.00    |        |        |        |        |      |
| 3. Movement time        | 0.07  | -0.09   | 1.00   |        |        |        |      |
| 4. Total response time  | 0.07  | 0.01    | 0.99** | 1.00   |        |        |      |
| 5. Maximum deviation    | 0.12  | -0.25*  | 0.31** | 0.29** | 1.00   |        |      |
| 6. Area under the curve | 0.21* | -0.23*  | 0.27** | 0.24** | 0.95** | 1.00   |      |
| 7. X-flips              | 0.06  | -0.23** | 0.06   | 0.04   | 0.46** | 0.46** | 1.00 |

Note. \* $p < .05$ ; \*\* $p < .01$ . Condition: Faking = 1 and Honest = 0.  $n = 68$  in both conditions.

FIGURE 4.

Mean mouse trajectories by condition



tracking can help with the evaluation of item pairs as the item pairs that are equal in social desirability should lead to greater areas under the curve. In operational settings, the area under the curve and other mouse tracking metrics may hold the potential for identifying those who are faking. With a sufficiently large data base, normative distributions for each of these metrics can be constructed. Individuals who are outliers in these distributions can be flagged as potential fakers. Unlike other methods to detect faking such as social desirability scales, no additional items are needed with mouse tracking, and the method is unobtrusive. Clearly more research using mouse tracking is needed before implementing the methodology in operational settings, but the growing use of forced-choice measures provides an opportunity to conduct this research.

#### Limitations and Future Directions

A primary limitation of this study is that select items, rather than an intact personality inventory, were used. Choices about item inclusion and pairings were based solely on social desirability ratings and did not take into

account equal representation of the five personality factors. This was a deliberate choice as this study was a starting point in the investigation of using mouse tracking to study cognitive processing during faking. Although conclusions about individual personality factors cannot be made here, there is the possibility to do so in future studies by including full representation of items from each of the five personality factors and creating item pairs based on factor as well as desirability. This will allow for an understanding of how personality factor influences how forced-choice item pairs are cognitively processed when they are equal or unequal in social desirability.

As was found in this study, the use of these behavioral methodologies may require presenting a greater number of item pairs. Several of the analyses found moderately sized effects, but they were not statistically significant due to the larger standard errors that are common for behavioral research methods. To minimize standard error within-person, a greater number of item pairs are needed.

Research using mouse tracking has predominately been conducted using computers with a traditional computer

mouse. At this point, it is not known if the findings from a traditional computer mouse will be replicated using trackpads such as those on laptop computers or touch screens, which are common on mobile devices and many computers. Trackpads and touch screens may function differently than a traditional mouse for capturing motor movements associated with cognitive processing. Future research should examine the use of mouse tracking on trackpads and touch screens. Also, more research examining the use of mouse tracking with individuals who are left handed is needed. The research to date has focused primarily on right-handed individuals or has not systemically considered the impact of left-hand dominance.

Although most of the literature on cognitive processes underlying faking has used a directed faking approach, this approach does have potential limitations including the generalizability of the findings to operational settings. Future research is needed that examines the usefulness of mouse tracking in field settings to determine the generalizability of the findings of this study. Research should also examine other types of faking instructions that may come closer to operational faking such as asking participants to fake as if they were applying to a specific job (e.g., customer service representative).

## REFERENCES

- Cao, M., & Drasgow, F. (2019). Does forcing reduce faking? A meta-analytic review of forced-choice personality measures in high-stakes situations. *Journal of Applied Psychology, 104*(11), 1347-1368.
- Dilchert, S., & Ones, D. S. (2012). Application of preventive strategies. In M. Ziegler, C. MacCann, & R. Roberts (Eds.) *New perspectives on faking in personality assessment* (pp. 177-200). Oxford University Press.
- Duran, N. D., Dale, R., & McNamara, D. S. (2010). The action dynamics of overcoming the truth. *Psychonomic Bulletin & Review, 17*(4), 486-491. <https://doi.org/10.3758/PBR.17.4.486>
- Ellingson, J. E., & McFarland, L. A. (2011). Understanding faking behavior through the lens of motivation: An application of VIE theory. *Human Performance, 24*(4), 322-337.
- Fine, S., & Pirak, M. (2016). Faking fast and slow: Within-person response time latencies for measuring faking in personnel testing. *Journal of Business & Psychology, 31*(1), 51-64.
- Freeman, J. B. (2009). *MouseTracker* (Version 2.83).
- Freeman, J. B., & Ambady, N. (2010). *MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. Behavior Research Methods, 42*(1), 226-241. <https://doi.org/10.3758/BRM.42.1.226>
- Freeman, J. B., Ambady, N., Rule, N. O., & Johnson, K. L. (2008). Will a category cue attract you? Motor output reveals dynamic competition across person construal. *Journal of Experimental Psychology: General, 137*(4), 673-690. <https://doi.org/10.1037/a0013875>
- Goffin, R. D., & Boyd, A. C. (2009). Faking and personality assessment in personnel selection: Advancing models of faking. *Canadian Psychology/Psychologie canadienne, 50*(3), 151.
- Griffith, R. L., Lee, L. M., Peterson, M. H., & Zickar, M. J. (2011). First dates and little white lies: A trait contract classification theory of applicant faking behavior. *Human Performance, 24*(4), 338-357.
- Hehman, E., Stolier, R. M., & Freeman, J. B. (2015). Advanced mouse-tracking analytic techniques for enhancing psychological science. *Group Processes & Intergroup Relations, 18*(3), 384-401.
- Holden, R. R., Kroner, D. G., Fekken, G. C., & Popham, S. M. (1992). A model of personality test item response dissimulation. *Journal of Personality and Social Psychology, 63*, 272-279.
- Komar, S., Komar, J. A., Robie, C., & Taggar, S. (2010). Speeding personality measures to reduce faking: A self-regulatory model. *Journal of Personnel Psychology, 9*(3), 126.
- König, C. J., Merz, A.-S., & Trauffer, N. (2012). What is in applicants' minds when they fill out a personality test? Insights from a qualitative study. *International Journal of Selection and Assessment, 20*(4), 442-452.
- Kuncel, N. R., & Tellegen, A. (2009). A conceptual and empirical reexamination of the measurement of the social desirability of items: Implications for detecting desirable response style and scale development. *Personnel Psychology, 62*(2), 201-228.
- Kuncel, N. R., Goldberg, L. R., & Kiger, T. (2011). A plea for process in personality prevarication. *Human Performance, 24*(4), 373-378.
- Meißner, M., & Oll, J. (2018). The promise of eye-tracking methodology in organizational research: A taxonomy, review, and future avenues. *Organizational Research Methods, 22*(2), 590-617. doi:10.1177/1094428117744882
- Mesmer-Magnus, J., & Viswesvaran, C. (2006). Assessing response distortion in personality tests: A review of research designs and analytical strategies. In R. Griffith & M. Peterson (Eds.), *A closer look at applicant faking behavior* (pp. 85-114). Information Age.
- Monaro, M., Gamberini, L., & Sartori, G. (2017). The detection of faked identity using unexpected questions and mouse dynamics. *PLoS One, 12*(5), e0177851.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Are we getting fooled again? Coming to terms with limitations in the use of personality tests for personnel selection. *Personnel Psychology, 60*(4), 1029-1049.
- Robie, C., Brown, D. J., & Beaty, J. C. (2007). Do people fake on personality inventories? A verbal protocol analysis. *Journal of Business and Psychology, 21*(4), 489-509. <https://doi.org/10.1007/s10869-007-9038-9>
- Scherbaum, C. A., Sabet, J., Kern, M. J., & Agnello, P. (2013). Examining faking on personality inventories using unfolding item response theory models. *Journal of Personality Assessment, 95*(2), 207-216. <https://doi.org/10.1080/00223891.2012.725439>
- Shoss, M. K., & Strube, M. J. (2011). How do you fake a personality test? An investigation of cognitive models of impression-managed responding. *Organizational Behavior and Human Decision Processes, 116*(1), 163-171.
- Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous at-

- traction toward phonological competitors. *Proceedings of the National Academy of Sciences*, 102(29), 10393–10398. <https://doi.org/10.1073/pnas.0503903102>
- van Hooft, E. A. J., & Born, M. Ph. (2012). Intentional response distortion on personality tests: Using eye-tracking to understand response processes when faking. *Journal of Applied Psychology*, 97(2), 301–316.
- Vasilopoulos, N. L., Reilly, R. R., & Leaman, J. A. (2000). The influence of job familiarity and impression management on self-report measure scale scores and response latencies. *Journal of Applied Psychology*, 85(1), 50–64.
- Zhang, B., Sun, T., Drasgow, F., Chernyshenko, O. S., Nye, C. D., Stark, S., & White, L. A. (2020). Though forced, still valid: Psychometric equivalence of forced-choice and single-statement measures. *Organizational Research Methods*, 23(3), 569–590.
- Ziegler, M. (2011). Applicant faking: A look into the black box. *TIP: The Industrial-Organizational Psychologist*, 49(1).
- Ziegler, M. E., MacCann, C. E., & Roberts, R. D. (2012). Faking: Knowns, unknowns, and points of contention. In M. Ziegler, C. MacCann, & R. Roberts (Eds.) *New perspectives on faking in personality assessment* (pp. 3–16). Oxford University Press.

RECEIVED 10/01/20 ACCEPTED 04/08/21