UNIVERSITY OF ALBERTA
FACULTY OF ARTS
Department of Economics

Working Paper No. 2023-12

Whoever You Want Me to Be: Personality and Incentives

Andrew McGee
University of Alberta

Peter McGee
University of Arkansas

December 2023

Whoever you want me to be: personality and incentives[*]

Andrew McGee[†] Peter McGee[‡]

December 2023

Abstract:

What can employers learn from personality tests when applicants have incentives to misrepresent themselves? Using a within-subject, laboratory experiment, we compare personality measures with and without incentives for misrepresentation. Incentivized personality measures are weakly to moderately correlated with non-incentivized measures in all treatments. When test-takers are given a job ad indicating that an extrovert (introvert) is desired, extroversion measures are positively (negatively) correlated with IQ. Among other characteristics, only locus of control appears related to faking on personality measures. Our findings highlight the identification challenges in measuring personality and the potential for correlations between incentivized personality measures and other traits.

Keywords: personality, measurement, hiring, screening, experiments

JEL codes: C91, D82, M50

[*] The authors thank the editor, Glen Waddell, an anonymous referee, participants at annual meetings of the Canadian Economics Association, the Economic Science Association, the European Association of Labour Economists, the Society of Labor Economics, and the IZA World Labor Conference and seminar audiences at Simon Fraser, the University of Arkansas, and the University of Alabama for their helpful comments. Catherine Michaud-Leclerc and Hanh Tong provided outstanding assistance. We gratefully acknowledge the financial support from SSHRC Insight Development Grant #430-2013-00700. Research Ethics approval was received from Simon Fraser University (2012s0590). All remaining errors are our own.

Data Availability Statement: The data that support the findings of this study are openly available in OpenICPSR at https://doi.org/10.3886/E194887V1, reference number openicpsr-194887.

[†] Corresponding Author: Department of Economics, University of Alberta, 8-14 Tory Building, Edmonton, AB T6G 2H4, Canada, and IZA. E-mail: mcgee1@ualberta.ca.

[‡] Department of Economics, University of Arkansas, Business Building 402, Fayetteville, AR 72701. Email: pmcgee@walton.uark.edu.

1. Introduction

Using personality tests to screen job applicants is associated with higher productivity at the firm level (Ichniowski et. al., 1997) and more productive hires with lower turnover (Autor and Scarborough, 2008; Hoffman et al., 2017). In the United States, 60-70% of employers use online personality tests to screen applicants (Weber and Dwoskin, 2014), but what employers learn from these tests is unclear as applicants have incentives to misrepresent themselves. When applying for a job, what applicant strongly agrees with the statement (from the online test of a major retailer), "When I encounter very difficult problems, I tend to move on to something else"? Surely some applicants lie—potentially a lot. Personality testing affects the careers of millions of workers each year through the tests' influence on employers' interviewing, hiring, and promotion decisions. Understanding the information conveyed by incentivized personality tests to employers is critical for assessing the likely impact of these tests on different workers.

We investigate what can be learned from personality tests when individuals have incentives to misrepresent themselves by conducting a within-subject, laboratory experiment addressing three questions. First, to what extent are incentivized personality scores correlated with non-incentivized scores for the same individuals? Second, are there non-personality traits that influence incentivized personality scores because they either make individuals better at misrepresenting themselves or more inclined to do so? Third, how does information about desired personalities influence incentivized scores and their relationships to other traits?

In the first "Baseline" session of the experiment, subjects were paid $7 to complete a questionnaire measuring, among other things, the Big Five personality traits. No further incentives were provided. About a week later, subjects participated in the second "Evaluation" session. Subjects were given a job ad and told they would complete personality and IQ tests after

which a $25 bonus would be awarded to the subject who best met the "hiring" criteria implicit in the ad.

The treatments differed in the ads presented to subjects. In the *Extroversion* treatment, the ad indicated that an extroverted worker was desired. We prime a single trait because Hough and Oswald (2000) report that employers use personality tests to identify extroverts.[1] The ad in the *Introversion* treatment indicated that an introverted worker was desired. In the *No Priming* treatment, the ad provided little information about desired personality traits.

We present three findings from the experiment—in which subjects misrepresent themselves significantly on all five personality traits in all treatments.[2] First, incentivized personality trait measures are only weakly to moderately correlated with non-incentivized measures with correlations ranging between 0.2 and 0.6. Furthermore, the information about desired personality traits available to subjects has little effect on these correlations between incentivized and non-incentivized trait measures. By contrast, retest correlations for individuals taking the Big Five test a week or a month apart without incentives are between 0.8 and 0.9.

Second, we find little evidence that the non-Big Five characteristics measured in our experiment are related to incentivized personality measures in all treatments. When individual characteristics influence either the ability to fake or willingness to do so, these characteristics should be correlated with incentivized personality measures whenever faking occurs, but none of the measured characteristics in our experiment are correlated with any incentivized trait measure

[1] McGee and McGee (2020) report that employers in the UK were most likely to use personality tests when filling vacancies for managers and sales workers—occupations requiring significant interaction with others.
[2] By misrepresentation, we mean the difference between scores in the Baseline and Evaluation sessions. The Baseline scores are themselves noisy measures of the underlying traits, but we interpret deviations from these scores in the primed or socially desirable direction as evidence of misrepresentation.

in all treatments. Indeed, most of the measured characteristics in the experiment are uncorrelated with incentivized personality scores in all treatments.

Third, we find that the information about desired personality traits available to subjects influences the correlations between incentivized scores and IQ and locus of control beliefs. In particular, we find that more intelligent individuals are better at adapting their extroversion responses when information about desired personalities is provided. In the *Extroversion* (*Introversion*) treatment, a one standard deviation increase in IQ is associated with a 0.17 SD (0.19 SD) increase (decrease) in the incentivized extroversion score, but IQ is uncorrelated with incentivized extroversion in the *No Priming* treatment when there is no information to extract. Likewise, we find little evidence that IQ is correlated with incentivized measures of the personality traits that were not primed. The correlations between IQ and the incentivized extroversion measures imply that subjects differ in their ability to extract information about desired traits from ads, which could explain why this information has relatively little effect on the correlations between the incentivized and non-incentivized trait measures.

The correlations between locus of control beliefs and incentivized personality measures—primarily conscientiousness and neuroticism—also appear to be moderated by the information available as these beliefs are uncorrelated with incentivized personality measures in the *Extroversion* treatment. In this treatment, the priming is most explicit and all subjects—regardless of their locus of control beliefs—may understand that faking effort will be rewarded. By contrast, in the *Introversion* and *No Priming* treatments subjects may have been less certain about the return to faking effort—thus opening a channel through which beliefs about the role of one's effort in determining outcomes may influence test-taking behavior.

Our findings underscore the identification challenges in measuring traits discussed in Almlund et al. (2011)—namely the need to account for incentives and situations when measuring personality. To date, research on the role of incentives largely focuses on measures of cognitive ability in environments without strong incentives, which have been shown to be correlated with non-cognitive traits (e.g., Borghans et al., 2008; Duckworth et al., 2011; Segal, 2012; Heckman and Kautz, 2012). An exception is Chen et al. (2020), who document that non-pecuniary (but not pecuniary) incentives for performance on IQ tests influence accompanying personality measures. Our findings suggest that accounting for incentives is of first-order importance when measuring personality, while accounting for the information available—though not without consequence—is less important.

Relative to the psychology literature on faking discussed in the next section, our primary contribution is the use of a within-subject experiment with incentives but not directives for faking that allows us to investigate the correlates of faking behavior and the moderating effects of information about desired candidates on these relationships. We demonstrate that IQ and locus of control beliefs can be correlated with incentivized personality measures depending on the information about desired traits available to test-takers. This finding is both novel to the psychology literature and potentially very important for understanding the success of personality testing as a screening mechanism.

As for our findings' implications for research on the role of personality in labor markets, researchers using personality measures from incentivized environments (e.g., Dal Bo et al., 2013) can take heart in the fact that these measures are likely correlated with non-incentivized measures even if the correlations may be modest. At the same time, our findings imply the

potential for omitted variables biases when employing incentivized personality measures as controls given that these measures may be correlated with IQ and locus of control beliefs.

For firms using personality testing in hiring, our findings are a mixed bag. Incentivized personality tests are noisy measures of personality traits if their correlations with non-incentivized measures are anything to go by, and firms should appreciate the limits of using such tests to screen applicants. At the same time, incentivized personality measures may help firms learn indirectly about traits (IQ and locus of control beliefs) that influence productivity. In the experiment, for instance, eliminating half of subjects based on incentivized extroversion scores would do nearly as well at eliminating low IQ subjects as it would at eliminating subjects based on their non-incentivized extroversion scores in the *Extroversion* and *Introversion* treatments. While easily administered and effective general mental ability tests exist (Schmidt and Hunter [1998]), employers in the United States may be wary of giving such tests if doing so exposes them to litigation risks in light of the *Griggs v. Duke Power Co.* (401 U.S. 424, 1971) decision.[3] Enabling employers to aggregate information about applicants on many dimensions with a single test may be one reason why such testing is successful and popular with employers.

For job-seekers, our findings are also a mixed bag. Evidence from job ads indicates that firms most often demand extroversion, conscientiousness, and openness (Brenčič and McGee, 2023a), while surveys of hiring managers find that conscientiousness is the trait most valued by employers (Dunn et al., 1995; Wehner et al., 2022). Our study confirms that incentivized trait measures do convey information about non-incentivized trait measures, and thus employers can use the tests to eliminate applicants with undesired traits. Individuals with traits such as introversion are likely to be disadvantaged when employers use personality tests to screen out

[3] The *Griggs v. Duke Power Co.* (401 U.S. 424) decision does not prohibit the use of IQ tests in hiring, but rather establishes that employers must show that such tests are necessary.

applicants with these traits. Such disadvantages, however, could be modest depending on the individual's ability to fake. Further, more intelligent applicants may be able to compensate for undesired traits if employers convey information about the traits sought in employees.

Concerning the external validity of our findings, job applicants may have incentives not present in our experiment that reduce the attractiveness of faking. Specifically, applicants may wish to avoid being caught faking or working in a job for which they are a bad personality "fit." That said, there are reasons to suspect that these incentive may be of limited importance in practice. Online job tests taken at the time of application are processed by algorithms that often reduce the tests to a single score predicting the probability of an applicant's success as an employee (Raghavan et al., 2020).[4] Applicants have little to fear from misrepresentation in the way of reputational concerns given that hiring managers typically do not see applicants' detailed responses.[5] Furthermore, young workers readily leave jobs that are poor fits as they learn about jobs (e.g., Gielen, 2013). Reputational concerns and concerns about fit are likely less important to applicants than progressing to the interview stage of hiring. Thus actual job applicants solve an optimization problem—passing a first screening hurdle to advance to the next stage of the hiring process—similar to that solved by the subjects in our experiment despite the absence of these incentives in our experiment.

2. Personality testing and hiring

Industrial/organizational psychologists have long known that individuals can "fake" (e.g., Velicer and Weiner, 1975; Kroger and Wood, 1993; Holden and Hibbs, 1995) and debated the

[4] In Hoffman et al. (2017), for example, the screening algorithm produced a simple "green-yellow-red" score advising hiring managers whether to consider the applicant.

[5] Moreover, job seekers encounter these tests frequently and likely understand that the tests are used to winnow applicant pools rather than to select individuals for hiring. Among subjects in the experiment who had been employed, 45% had taken such tests as part of a job application.

usefulness of personality tests in hiring (e.g., Ghiselli and Barthol, 1953).[6] This literature established that test-takers can fake both "good" and "bad" responses (Hough et al., 1990; Furnham, 1997), find it equally easy to fake all Big Five personality traits (Viswesvaran and Ones, 1999), and tilt their answers toward socially desirable responses (Ones and Viswesvaran, 1998; Paulhus, 2002; Donovan et al., 2003). Like these papers, we also find that our subjects distort their responses significantly in the socially desirable directions on all Big Five traits regardless of the experimental priming.

The faking research can be divided into two sets of studies. In "fake good" studies, subjects are asked to misrepresent themselves in (usually) favorable ways (e.g., Velicer and Weiner, 1975; Furnham, 1997; McFarland and Ryan, 2000). The second set of studies compares the personality scores of incumbents to applicants for the same position (e.g., Hough et al., 1990; Becker and Colquitt, 1992; Rosse et al., 1998; Birkeland et al., 2006). The "fake good" studies tend to find larger response distortions than applicant-incumbent studies, but these studies typically explicitly direct subjects to fake without providing incentives to do so.[7]

Our contribution to this literature is in part methodological. Unlike the between-subject applicant-incumbent studies, our within-subject study allows us to assess the relationships between incentivized personality measures and non-incentivized personality measures and other characteristics. Rosse et al. (1998) note that the job applicants in their applicant-incumbent sample exhibited substantial heterogeneity in response distortion as measured by impression management scores with some engaging in extreme response distortion and others engaging in none. We observe similar heterogeneity in faking when comparing incentivized to non-

[6] Excellent summaries of the debates surrounding personality tests in selection can be found in Rothstein and Goffin (2006), Morgeson et al. (2007a,b), Ones et al. (2007), and Hough and Oswald (2008).
[7] Exceptions include Dwight and Donovan (2003), Vasilopoulos et al. (2005), and Huber et al. (2021) who used between-subject, "fake good" designs in which the top-performing subjects were paid bonuses as in our experiment.

incentivized scores, and a primary objective of the study was to determine whether this heterogeneity could be explained by other individual characteristics.

Unlike "fake good" studies without incentives, our study examines how misrepresentation in response to incentives but not explicit directives for faking—the testing conditions faced by most job applicants—affects the ranking of applicants by personality traits—an issue around which there is debate (e.g., Morgeson et al., 2007 a, b; Ones et al., 2007). Moreover, our study considers how the provision of information about desired personalities in job ads affects incentivized personality measures and their relationships to other characteristics. In qualitative studies of faking behavior, Ziegler (2011) finds that respondents consider the importance of test items in terms of situational demands when deciding whether to fake, while König et al. (2012) report that the majority of applicants interviewed reference the expectations of the organizations to which they applied when describing their faking strategies. Krammer (2020) finds that the perceived relevance of test items for one's desired profession influenced faking on incentivized personality tests taken in the college admissions process. Our treatment manipulation varies the information subjects see in job ads given that many applicants are routed to job tests directly from ads and the information in these ads undoubtedly influences applicants' understanding of situational demands. By contrast, "fake good" studies using job titles to direct subjects to fake as they think workers in occupations would (e.g., Velicer and Weiner, 1975; Furnham, 1990) rely on stereotypes rather than providing information about desired traits.

The study most similar to ours is that of Tett et al. (2012), who use a within-subject, "fake good" design without incentives in which subjects were given job descriptions and asked to respond "in a way that will make you appear as an ideal job candidate."[8] They find that

[8] Tett el al. did not explicitly prime particular traits.

generalized intelligence explains some observed faking (as evidenced by changes in R^2), but their study was not designed to establish conditions under which intelligence affects incentivized personality measures. By contrast, our study indicates that IQ influences incentivized personality measures when information about desired traits is provided and only for those traits being sought. Moreover, while Tett et al. speculate that traits such as Machiavellianism and dishonesty may influence faking, our study measures these traits and tests these hypotheses.[9]

An ideal experiment would utilize a within-subject design in a hiring context so that the incentives of test-takers are those of actual applicants, but the few studies using this approach highlight the attendant difficulties.[10] In Griffith et al. (2007) and Isaacson et al. (2009), for example, only 43% (n=60) and 6% (n=196) of job applicants, respectively, responded to researchers' non-incentivized surveys resulting in small and potentially non-representative samples. Ellingson et al. (2007) used the database of a personality testing firm to identify 218 individuals who took tests in a non-incentivized test development setting and when applying for jobs, but these individuals were applying for different jobs and the tests were taken as much as seven years apart. Boyce (2005) found that beliefs about whether others fake were related to faking among individuals who applied to and worked at an amusement park. Successful applicants, however, tend to have favorable traits and less need to fake than other applicants. As a result, faking in such samples is less evident and the range of measures restricted.

In a meta-analysis of within-subject studies comparing personality measures taken in both high- and low-stakes settings, Hu and Connelly (2021) estimate correlations between the high-

[9] Kleinmann et al. (2011) speculate that the link between incentivized personality measures and productivity may result from applicants differing in their ability to identify the criteria used to evaluate responses if this ability is related to productivity, but they provide no evidence to this effect.
[10] Hu and Connelly (2021) identify 12 samples from published studies and 8 samples from unpublished studies employing within-subject designs in high- and low-stakes settings.

stakes and low-stakes Big Five measures ranging between 0.56 and 0.65—only slightly higher than our estimates ranging mostly between 0.4 and 0.6. These within-subject studies in hiring contexts, however, examined neither the effects of information nor the correlations between incentivized personality measures and the characteristics measured in our experiment.

Finally, our findings highlight the fact that estimates of the relationship between personality testing in hiring and employee and firm outcomes could be influenced by variation in the information about desired traits available to test-takers. This serves as a cautionary note for research on testing using samples from single firms (e.g., Autor and Scarborough, 2008; Hoffman et al., 2017). Given that firms may advertise positions with language that induces correlations between incentivized personality measures, IQ and locus of control, estimates of the benefits of testing in one firm may not generalize to other firms.

3. Modeling Behavior on Incentivized Personality Tests

In Almlund et al.'s (2011) Roy model of personality testing, personality tests consist of items P_j $(j = 1, \dots, J)$ with payoffs R_j for each item. Responses are assumed to be a function of characteristics (θ) and effort e_j:

$$P_j = \phi_j(\theta, e_j)$$

The cost of expending effort on item j is given by $C_j(\theta, e_j)$. Test-takers allocate effort to maximize the expected returns to performance given the available information $\mathbb{I}$:

$$\max_{\{e_j\}_{j=1}^{J}} E\left\{\sum_{j=1}^{J}[R_j\phi_j(\theta, e_j) - C_j(\theta, e_j)] \,\middle|\, \mathbb{I}(\theta)\right\}$$

The notion that some traits may be related to "faking" is captured by either the cost function or the production function ϕ_j. Specifically, the costs of manipulating performance through effort might be decreasing in some traits (i.e., $\frac{\partial^2 C_j(\theta, e_j)}{\partial e_j \partial \theta} < 0$). For instance, intelligent

individuals may find exerting effort less onerous. Alternatively, some individuals may dislike misrepresenting themselves less than others. The second example highlights that θ may include preference parameters as well as traits like personality and intelligence.[11]

Individuals maximize the expected returns given their information set $\mathbb{I}(\theta)$, where the uncertainty surrounds the returns to tasks R_j. The information set $\mathbb{I}(\theta)$ may depend on characteristics θ as well as the situation if, for instance, more intelligent or socially astute individuals are better at extracting cues from situations than others.

Almlund et al. observe that in order to identify personality traits from performance on test items P_j one needs to account for effort levels, incentives, situations, information sets, etc. For simplicity, assume that truthful reporting on item j implies that $e_j = 0$.[12] Thus "faking" involves the manipulation of performance through effort ($e_j > 0$). The model gives rise to the following research questions:

Research question #1: How important are incentives when measuring personality?

We obtain personality measures from the same subjects in environments with and without incentives for misrepresentation. Assessing the correlation between the resulting personality measures sheds light on the empirical importance of incentives when measuring personality.

Research question #2: Are incentivized personality measures influenced by traits other than personality?

In the model, the ability to fake or the costs of faking may depend on traits other than personality. As a consequence, incentivized personality measures may be correlated with trait

[11] As Almlund et al. note, allowing the costs of effort to vary with θ is empirically indistinguishable from the assumption that traits θ and effort e_j are complements or substitutes in the performance of task j (i.e., $\frac{\partial^2 \phi_j(\theta, e_j)}{\partial e_j \partial \theta} \gtrless 0$).

[12] If test-takers have no incentive to misrepresent themselves, one might assume that $R_j = 0$ for all j, and subjects maximize their returns by minimizing effort costs and set $e_j = 0$ for all j.

measures with which they are uncorrelated in the absence of faking. A subtler point is that if a correlation arises between a personality measure and another characteristic because the latter makes faking more productive or less costly, then a correlation should exist in *all* situations that induce faking.[13]

We hypothesized that more intelligent individuals might be better at adapting responses toward socially desirable responses while characteristics including Machiavellianism, self-monitoring, self-deception, impression management, and views about the acceptability of lying might be correlated with one's willingness to engage in misrepresentation. We hypothesized that optimistic subjects might feel less need to misrepresent themselves if their optimism extended to beliefs about their performance and that risk averse subjects might be more willing to lie in order to mitigate earnings variance.[14] Finally, we hypothesized that individuals with an internal locus of control who believe that outcomes are primarily a function of their own efforts might believe the return to faking to be higher and exert more effort misrepresenting themselves.

Research question #3: How does the information available to test-takers influence the relationships between measured personality and other traits?

Traits such as intelligence may influence the information individuals extract from environments and thus expectations about the returns to performance. As such, whether correlations exist between a measured characteristic and incentivized personality measures may depend on the information available.[15]

---

[13] Consider two uncorrelated traits, θ_1 and θ_2, and a test measuring θ_1 consisting of a single item P_1. Let performance on the item be given by $P_1 = \alpha_1\theta_1 + \alpha_2\theta_2 + \alpha_3 e_1 + \alpha_4 e_1\theta_2$. If $\alpha_2 = 0$, then P_1 will be uncorrelated with θ_2 in the absence of faking (i.e., when $e_1 = 0$). When a complementarity between effort e_1 and trait θ_2 exists (i.e., $\alpha_4 \neq 0$), the personality measure P_1 will be correlated with θ_2 even if $\alpha_2 = 0$ whenever faking occurs (i.e., $e_1 > 0$).

[14] A role for risk preferences could be introduced by relaxing our assumption that test-takers maximize a linear payoff function.

[15] Suppose $P_1 = \alpha_1\theta_1 + \alpha_3 e_1$. That is, trait θ_2 affects the personality measure P_1 neither directly (because $\alpha_2 = 0$) nor through faking ability (because $\alpha_4 = 0$). If the return to task performance R_1 is known or all subjects have the

4. Experimental Design

4.1 Baseline Session

In the Baseline session, subjects answered a questionnaire consisting of a Big Five assessment (DeYoung et al., 2007), the Balanced Inventory of Desirable Responding (Paulhus, 1984), a self-monitoring scale (Snyder, 1974), a Machiavellianism scale (Jackson, 1994), an optimism-pessimism scale (Scheier et al., 1994), questions about the acceptability of lying with different motivations (McLeod and Genereux, 2008) and demographic questions.[16,17,18] Subjects were informed at the outset that they would only receive a $7 payment for participation.

The Baseline session lasted 15 to 20 minutes for most subjects. The instructions informed subjects that their participation made them eligible for a future experiment. After paying each subject, our research assistant offered to sign the subject up for an Evaluation session the following week. In most cases, subjects agreed, but subjects were not required to participate in the additional session and not all did.[19] As a result, most subjects participated in the Baseline and

---

same expectations concerning this return because no information is available, then personality measure $P_1 = \alpha_1\theta_1 + \alpha_3 e_1^*(R_1)$ will be uncorrelated with θ_2 where $e_1^*(R_1)$ is the optimal effort on task 1 given the known or common expected return to item 1, R_1. If instead R_1 is unknown and θ_2 affects individuals' inference concerning the return, then $P_1 = \alpha_1\theta_1 + \alpha_3 e_1^*(E[R_1|\theta_2])$ and P_1 may be correlated with θ_2 even if θ_2 has no effect on the ability to fake or costs of faking.

[16] DeYoung et al.'s (2007) 100-item Big Five assessment asks subjects to indicate how well a statement describes them using a 5-point Likert scale ranging from "Strongly disagree" to "Strongly agree." For instance, "I start conversations" is an item on the extroversion scale. Summing up the 20 items associated with each trait accounting for reverse scoring leads to trait scores between 20 and 100. Comparing this test to those of two major retailers using online tests to screen applicants, we found that 90% and 46% of the items on the retailers' tests were identical or nearly identical to items on the DeYoung et al. test. Given that the employers' tests also include items that are clearly not related to personality, the similarities suggest that the DeYoung et al. test is a reasonable approximation for the personality component of tests used by actual employers.

[17] The experiment was conducted using z-Tree (Fischbacher, 2007).

[18] We elected to use the non-incentivized survey measures of the acceptability of lying given that we test hypotheses about correlates of faking at the individual (rather than group) level and that preferences for honesty may depend on the context and motivation. Incentivized measures of preferences for honesty exist, but these either can detect lying only at the group level (e.g., Fischbacher and Föllmi-Heusi, 2013) or conflate preferences for honesty with strategic considerations (see Abeler et al. (2019) footnote 2 for a discussion).

[19] Subjects were not obliged to participate in the Evaluation session, but 80% of subjects did so. Appendix Table 1 reports summary statistics for the characteristics measured in the Baseline for non-returnees and returnees in our sample; there are no statistically significant differences between the two groups.

Evaluation sessions one week apart, but nothing prevented subjects from registering for the Evaluation session more than one week after the Baseline.

4.2 Evaluation Session

In the Evaluation session, subjects were placed in groups of four subjects. Subjects did not interact with members of their group and were not aware which subjects were in their group. The instructions informed subjects that they would receive a job description and then complete personality and intelligence tests and that a $25 bonus would be awarded to one member of each group on the basis of the tests. The job description varied across treatments. The instructions, survey instruments and job descriptions for each treatment are provided in the Appendix.

After the instructions, subjects had 20 minutes to complete Raven's Progressive Matrices (RPM; Raven, 1998), a fluid intelligence test. They then had as much time as they needed to answer the same Big Five battery completed in the Baseline session (albeit with the items in a different order). Subjects learned whether they had received the bonus before completing the incentivized Holt and Laury (2002) risk aversion measure. Finally, subjects completed a short questionnaire measuring self-deception, impression management and locus of control.[20] Subjects were also asked whether they were employed, how many jobs they had worked, how many years of work experience they had, how many times they had taken a personality test as part of a job application, how confident they were that they knew what employers were looking for when looking for a job, and how the job description had affected their behavior in the experiment. In addition to a $7 participation fee and (if applicable) the $25 bonus, subjects were paid $0.20 for

[20] We elected to have subjects complete the locus of control battery (without incentives) in the Evaluation session to keep the Baseline session shorter in the hope that subjects would be more likely to return for the Evaluation session if hourly wages in the Baseline were higher. This design decision was informed by Doyle et al. (1977), who showed that locus of control measures are not affected by experimental interventions manipulating subjects' control over outcomes. We find no evidence in Table 1 that the locus of control measures differ across treatments.

each correct answer on the RPM and their earnings from one randomly selected realization of a paired lottery on the Holt-Laury instrument.

4.3 Treatments

Extroversion treatment*:* The *Extroversion* instructions informed subjects that "All members of a group will receive the same job ad. We will then administer an intelligence test and a personality test to determine who to "hire" for the job. We will weigh these two tests according to some criteria and one subject within each group who best meets these criteria will receive a bonus of $25; the remaining subjects in each group will not receive any bonus." The job ad indicated that a staffing firm was looking to hire an extrovert for a "client services representative" position. To indicate that an extrovert was desired, the job description incorporated words and phrases associated with extroversion such as "outgoing," "able to take initiative, be assertive," and "proficient at building and maintaining relationships" while attempting to avoid indicating that the firm was seeking workers with other traits.

Introversion treatment*:* The *Introversion* instructions were the same as in the *Extroversion* treatment, but the job description indicated that the staffing firm was looking for an introverted "client services representative." Specifically, the job description indicated that a "low-key" firm was looking for a "contemplative, reserved, independent" individual who "enjoys tackling projects solo" and who gets "the job done without making a splash or interrupting clients' normal business."

No Priming treatment*:* The instructions in the *No Priming* treatment were again the same as in the *Extroversion* treatment. The job ad in the *No Priming* treatment was also for a "client services representative" at a staffing firm, but the job description otherwise contained very little detail. The description indicated requirements such as "proficient in Microsoft Office" and "has

financial acumen"—requirements that were also in the *Extroversion* and *Introversion* job descriptions. The position title and tasks may themselves have "primed" subjects based on prior beliefs, but we attempted to provide as little additional priming as possible. [21]

The ads were based on an actual ad for a client services representative. This occupation was selected as the sort of job to which young, university graduates might apply. In the 2016 American Community Survey, 25% of respondents working in the services sales representative occupation were age 30 or less, 51% had less than a bachelor's degree, and 40% had bachelor's degrees (authors' calculations). As for the priming, Brenčič and McGee (2023a) find that 42% of job ads in this occupation signaled a demand for extroverts through terms associated with extroversion (nearly twice the proportion of ads expressing demands for any other trait), while less than 1% of such ads indicated a demand for introverts in a sample of job ads from Monster.com. At the same time, 36% of ads for this occupation used no terms associated with desired personality traits. As such, we view the *Extroversion* and *No Priming* treatments as highly representative of the sort of ads job seekers in this occupation might encounter.[22]

Retest sessions: DeYoung et al. (2007) report retest correlations for 90 subjects who were undergraduates at Canadian universities who completed the DeYoung et al. test twice without incentives one month apart. As a benchmark to establish the influence of incentives, we test the equality of the correlations between the incentivized and non-incentivized trait measures in the

[21] An additional treatment that did not involve variation in information about desired personalities is not discussed as it addressed a research question unrelated to the current study.

[22] A concern arising from using a laboratory experiment is that all subjects must complete the tests regardless of whether they would choose to apply for such a job. In the labor market, however, client services representative positions might only attract as applicants individuals with specific personality traits. While personality sorting in the labor market does occur, Brenčič and McGee (2023b) provide evidence that this sorting is very limited. In particular, occupations explain at most 8% of the variance in any personality trait among respondents in the National Longitudinal Survey of Youth 1997, and the occupation-specific means for each trait for 22 broad occupation categories differ from the means in the sample as a whole by less than 0.2 standard deviations in most cases. The latter implies that individuals with all sorts of personality traits are found in each and every occupation in the sample, meaning that applicants in most occupations are likely equally diverse in terms of their personalities.

treatments with the retest correlations reported by DeYoung et al. Because retest correlations can depend on how far apart in time subjects respond to the instrument, we also estimated the retest correlations at a one-week interval by having subjects at the University of Arkansas complete two Baseline sessions one-week apart, which we refer to as the *Retest* sessions.

Between 2014 and 2016, 474 subjects participated in both a Baseline and Evaluation session: 167 subjects in the *Extroversion* treatment, 160 in the *Introversion* treatment, and 147 in the *No Priming* treatment. Another 45 subjects participated in the *Retest* sessions.[23] Subjects were undergraduates at Simon Fraser University except for those in the *Retest* sessions.

Table 1 reports summary statistics for the non-Big Five measures by treatment, while Table 2 reports statistics for the Big Five measures in both the Baseline and Evaluation sessions. Using Kruskal-Wallis tests for differences across treatments in characteristics measured in the Baseline session, there are significant differences in openness (p=0.05), age (p=0.007), the fraction White (p=0.09), and willingness to lie to gain social acceptance (p=0.09).[24] All of our findings, however, persist when controlling for those traits that were unbalanced across treatments.

We also consider in Table 1 whether the treatments were balanced in terms of subjects' attentiveness in the Baseline session. We assume throughout the paper that subjects respond honestly in the Baseline such that their responses are indicative of the underlying traits measured

[23] The sample sizes in the treatments and *Retest* sessions were determined by power calculations. For the treatments, these calculations indicated that samples of 154 subjects were required to achieve a power of 0.8 to detect plausible, small correlations (i.e., 0.2) between measures using one-sided tests at the 5% significance level. The accidental double-counting of one session led us to fall slightly short of this target in the *No Priming* treatment, but our power in that treatment is still 0.78. Further power calculations suggest that at our sample sizes we have reasonable power (0.5) to detect correlations as small as 0.1 using one-sided tests at the 10% significance level. The sample size decision for the *Retest* sessions was guided by the desire to test the equality of the retest correlations with the correlations between the incentivized and non-incentivized personality measures in our treatments. These calculations indicated that 35 subjects were required to achieve power of 0.8 assuming samples in the treatments of 160, correlations between the incentivized and non-incentivized personality scores in the treatments of 0.5, and correlations in the retest sessions of 0.8 (i.e., roughly the correlations reported in DeYoung et al. (2007)).

[24] The statistically significant age difference across treatments is primarily the result of a single mature subject in the *Extroversion* treatment. Non-incentivized measures of the Big Five personality traits have been shown to be largely stable over time among adults (Cobb-Clark and Schurer, 2012).

with error, and thus misrepresentations reflect behavior in the Evaluation sessions. Variation in attentiveness and effort in the Baseline would complicate this interpretation.[25] To evaluate this possibility, we identified subjects who give the same response to every Big Five item on a screen more than once. Subjects were shown five Big Five items on the screen at any time and had to respond to each item before moving on to a new screen with five more items. Though it is entirely possible that subjects could have the same response for five consecutive items, doing so for several sets of items is more likely to reflect a lack of seriousness.[26] Only 32 subjects (6.8%) exhibited this potential inattention, and they were relatively evenly distributed across treatments. Moreover, the findings in the next section are robust to simply excluding these subjects.[27]

5. Findings

5.1 Main results

Figure 1 displays the kernel densities of the personality trait scores from the Baseline and Evaluation sessions by treatment. Subjects unambiguously responded to pecuniary incentives to misrepresent themselves—which is noteworthy insofar as we did not direct subjects to misrepresent themselves as in "fake good" studies. The p-values for the nonparametric Wilcoxon signed-rank tests of the equality of the mean trait scores in the Baseline and Evaluation sessions

[25] There are ex ante reasons to be skeptical that differences in Baseline effort influence our findings. First, the Big Five measures were designed for use in environments without incentives and have been repeatedly validated in these environments. The retest correlations at one-week and one-month intervals reported in our *Retest* sessions and in DeYoung et al. (2007) range between 0.8 and 0.9. These high correlations would be unlikely if subjects responded randomly or inattentively given that subjects were not incentivized to reproduce their earlier scores. Second, conditional on supplying effort, the least cognitively burdensome response strategy is likely to respond honestly in the absence of other incentives. Thus we focus on the possibility that some subjects reply randomly (i.e., supplying no or minimal effort).

[26] 127 subjects gave the same answer for all items on a screen at least once, 32 did so at least twice, and 13 did so three or more times. Only one subject did so more than 5 times.

[27] As an alternative measure of the seriousness with which subjects approached the Baseline session, we also identified 27 subjects (5.7%) who provided identical responses for at least one of four pairs of items in which the statements were nearly complete opposites. These subjects were also relatively evenly distributed across treatments, and our results are also robust to simply excluding these subjects.

in Table 2 are less than 0.001 for each trait in every treatment, and Kolmogorov-Smirnov tests reject the equality of the incentivized and non-incentivized distributions for every trait in all treatments at the 5% significance level. For every trait in all treatments, subjects misrepresent themselves on average in the socially desirable direction. In most cases, the Baseline and Evaluation means differ by a full standard deviation relative to the Baseline distribution. Even in the *Introversion* treatment, the incentivized extroversion scores are higher than the non-incentivized extroversion scores—though the increase is smaller and statistically different from the increases in the *Extroversion* and *No Priming* treatments.

Research question #1: How important are incentives when measuring personality?

Panel A of Table 3 reports the pairwise Pearson correlation coefficients between the Baseline and Evaluation scores for each of the Big Five personality traits by treatment. These correlations are weak to modest—particularly relative to the retest correlations in Column (4) at a one-month interval reported in DeYoung et al. (2007) and at a one-week interval in the *Retest* sessions. In the *Extroversion* treatment, for example, the correlation between the Baseline and Evaluation extroversion scores is just 0.38 compared to 0.83 in DeYoung et al. and 0.94 in the *Retest* sessions. Using z-tests, we reject the equality of all of the Pearson correlation coefficients in the treatments with the corresponding retest correlations reported in DeYoung et al. and in the *Retest* sessions at the 1% level.[28]

[28] As an alternative to z-tests, we can test the equality of correlations in Table 3 across treatments and with the correlations in the *Retest* sessions by regressing the standardized Evaluation session scores on the standardized Baseline scores, treatment indicators, and interactions between the treatment indicators and the Baseline scores. The coefficients of these interactions can be used for the appropriate hypothesis tests. The inferences using these tests of the equality of the correlations in the treatments with those in the *Retest* sessions are essentially the same as those reported in Table 3 except for openness in the *Introversion* treatment (p=0.16) and *No Priming* treatment (p=0.07). We use the z-tests because they require only the sample sizes and the correlations, which allows us to test the equality of the correlations in the treatments with those reported in DeYoung et al. (2007). To the extent that one might be concerned about the unbalanced nature of the samples for some characteristics in Table 1, one can include the measured characteristics from the Baseline session in these regressions to perform statistical tests on the partial

To assess whether the shifts in the distributions of incentivized scores preserve the ranking of subjects from most to least extroverted, Panels B and C report Spearman's rank correlation and Kendall's tau-b coefficients, respectively. Both statistics measure the extent to which the Baseline and Evaluation scores are related in a monotonic, order-preserving manner. For almost every trait in all treatments, the Spearman rank correlation and Kendall's tau-b are lower than the Pearson correlation coefficients. In the *Extroversion* treatment, the Spearman correlation between the Baseline and Evaluation extroversion scores is 0.31 compared to 0.93 in the *Test-Retest* sessions. Again using z-tests, we reject the equality of the statistics in Panels B and C with those in the *Retest* sessions at least at the 10% level for every trait except openness in the *Introversion* treatment.

The correlations have two implications. First, economists using incentivized personality measures should recognize that these are likely very noisy measures of the underlying trait if their correlations with non-incentivized measures are anything to go by. Standardizing personality measures for incentives appears to be very important. Second, firms using personality tests to trim the applicant pool (as Autor and Scarborough (2008) and Hoffman et al. (2018) suggest that firms do) should recognize the coarseness of this approach. To fix ideas, suppose firms used personality tests to eliminate applicants in the bottom half of the extroversion distribution, and assume that non-incentivized scores are error-free measures of the underlying traits. In the *Extroversion* treatment, only 51 of the 84 subjects in the bottom half of the non-incentivized extroversion distribution would be eliminated based on incentivized scores. By

correlations between the incentivized and non-incentivized scores across treatments and relative to the *Retest* sessions. The partial correlations and p-values from these tests are very similar to those reported in Table 3.

contrast, eliminating the bottom half of subjects when ranking them randomly would eliminate on average 42 of the 84 subjects in the bottom half of the non-incentivized distribution.[29]

Research question #2: Are incentivized personality measures influenced by traits other than personality?

Figure 2 displays the scatterplots by treatment for each Big Five trait of the Baseline scores against the Evaluation scores. The scatterplots reveal considerable heterogeneity in the extent to which subjects misrepresent themselves (where departures from the 45-degree line in the scatterplots are indicative of misrepresentations). Some of this heterogeneity may result from differences in characteristics that influence subjects' ability to fake or the costs of faking.[30]

To evaluate whether other traits influence the incentivized extroversion scores, we regress for each treatment the standardized, incentivized extroversion score on the standardized, non-incentivized extroversion score and standardized measures of the characteristics discussed in Section 3. The non-incentivized and incentivized extroversion scores are standardized using the Baseline and Evaluation session extroversion scores, respectively, of subjects in the same treatment, while the other measures are standardized using scores for subjects in all treatments.

Columns (1) to (3) of Table 4a report the OLS coefficient estimates for the *Extroversion*, *Introversion*, and *No Priming* treatments, respectively, with heteroskedasticity-robust standard errors given in parentheses. Similar to the Pearson correlation coefficients in Table 3, the

[29] In the *Introversion* treatment, 53 of the 80 subjects in the *top* half of the non-incentivized extroversion distribution would be eliminated based on the incentivized scores. In the *No Priming* treatment, 53 of the 74 subjects in the bottom half of the non-incentivized extroversion distribution would be eliminated based on the incentivized scores.
[30] The scatterplots also highlight the significant heteroskedasticity evident in our data and its mechanical nature as subjects with high Baseline scores have little scope to increase their incentivized scores compared to subjects with low Baseline scores.

coefficients of the non-incentivized extroversion score range from 0.413 in the *Extroversion* treatment to 0.483 in the *Introversion* treatment.[31]

In Section 3, we noted that if characteristics influence the ability to fake or the costs of faking for a personality trait, then these characteristics should be correlated with incentivized measures in any treatment in which subjects misrepresent themselves. Column (4) reports the p-values from Wald tests of the joint significance of the coefficients of the characteristic in a row across treatments. We reject the null that the coefficients in the three treatments are jointly equal to zero only for non-incentivized extroversion and IQ. A one standard deviation increase in IQ is associated a 0.173 standard deviation increase in the incentivized extroversion score in the *Extroversion* treatment and a 0.194 standard deviation reduction in the *Introversion* treatment. Notably, IQ is not correlated with incentivized extroversion in the *No Priming* treatment as it should be if intelligent individuals are simply better at faking extroversion.

Risk aversion is nearly jointly significant across treatments with a one standard deviation increase in risk aversion associated with a 0.158 standard deviation increase in incentivized extroversion scores in the *Extroversion* treatment and smaller and less precisely estimated increases in the *Introversion* and *No Priming* treatments. Risk averse subjects may attempt to minimize the risk of not winning the bonus by faking more in socially desirable directions. The remaining characteristics are not significantly correlated with incentivized extroversion scores in any treatment, nor are any jointly significant across treatments. These characteristics do not appear to influence subjects' misrepresentations where extroversion is concerned.[32]

[31] The coefficients of non-incentivized extroversion would equal the Pearson correlation coefficients in the absence of other controls given how we standardize the extroversion scores.

[32] We measured willingness to lie in four domains, but none of these measures were correlated with incentivized personality scores. Likewise, we used principal component analysis to construct a common factor score from the willingness to lie measures as well as for the self-deception, impression management, and Machiavellianism measures, but these index variables were also uncorrelated with the incentivized scores. Boyce (2005) similarly

Our focus on incentivized extroversion scores in Table 4a is motivated by Research Question 3, but the hypothesis that characteristics may influence the ability to fake or the costs of faking applies equally to all of the Big Five traits. Tables 4b to 4e report coefficient estimates for similar regressions by treatment using incentivized conscientiousness, agreeableness, openness, and neuroticism as the dependent variables, respectively.[33] The coefficient of IQ is jointly significant across the *Extroversion*, *Introversion*, and *No Priming* treatments for a single trait: openness. Self-deception, willingness to lie for personal gain, Machiavellianism and optimism are each statistically insignificant in every regression and jointly insignificant across treatments for the four additional personality traits. Risk aversion is positively correlated with incentivized agreeableness in the *No Priming* treatment while impression management is negatively correlated with incentivized agreeableness in the *Introversion* treatment, but otherwise risk aversion and impression management are uncorrelated with the incentivized trait measures.

Only locus of control appears to be broadly correlated with incentivized personality measures. Specifically, the locus of control coefficients are jointly significant across treatments for the incentivized conscientiousness and neuroticism measures and nearly so for openness and agreeableness. While jointly significant across treatments, however, the locus of control coefficient estimates vary across treatments in ways that suggest an interaction with the information available to subjects, which we discuss below. [34]

found that perceptions that others think faking is acceptable, ethics against lying, and beliefs that faking is lying were unrelated to faking in his sample of incumbents who also took personality tests as applicants.

[33] The trait scores are bounded between 20 and 100, and the scatterplots in Figure 2 make it appear as though many subjects have censored incentivized scores. In principle, this censoring could lead to inconsistent estimates of the coefficients in Tables 4a to 4e. In fact, censoring of the incentivized scores is fairly uncommon. For example, the incentivized extroversion scores are right-censored for only six observations in the *Extroversion* treatment, none in the *Introversion* treatment, and one observation in the *No Priming* treatment. Furthermore, we obtain estimates similar to those in Tables 4a to 4e using the semi-parametric censored least absolute deviations (CLAD) estimator (Powell, 1984) given the heteroskedasticity evident in our data (estimates available from the authors).

[34] Tables 4a to 4e control only for those individual characteristics that we hypothesized might influence faking; age, race, and gender are omitted. While IQ and race are correlated in our data, controlling for age, race, and gender in

Research question #3: How does the information available to test-takers influence the relationships between measured personality and other traits?

We first consider whether the information available to test-takers influences the correlations between incentivized and non-incentivized personality measures by testing whether the Pearson correlation coefficients for the *Extroversion* and *Introversion* treatments equal those in *No Priming* treatment in Panel A of Table 3. The correlations in the *Extroversion* treatment are lower than in the *No Priming* treatment, and using z-tests we reject the equality of the openness correlations in the *Extroversion* treatment and the conscientiousness correlation in the *Introversion* treatment with the corresponding correlations in the *No Priming* treatment. Using Holm-Sidak corrected p-values to account for the multiple hypotheses tested in Panel A, however, we fail to reject the null that the correlations in the *Extroversion* and *Introversion* treatments are equal to those in the *No Priming* treatment for all traits. Thus while incentives for misrepresentation result in incentivized personality scores that are only weakly to moderately correlated with non-incentivized scores as evidenced by the *No Priming* treatment, information about desired personality traits has relatively little effect on these correlations.

This does not mean that the provision of information about desired personalities has no effect on incentivized personality measures. In Table 4a, we reject the null hypotheses that the IQ coefficients in the *Extroversion* and *Introversion* treatments are equal to the coefficient in the *No Priming* treatment at the 1% level. Indeed, we fail to reject the null that the IQ coefficient in the *Extroversion* treatment has the same magnitude as that in the *Introversion* treatment but with opposite sign. We view the fact that the correlation between IQ and incentivized extroversion

similar regressions leads to estimates that are virtually identical to those in Tables 4a to 4e (estimates available from the authors). For a discussion of the relationships between race, gender, and the incentivized personality scores in the experiment, see McGee and McGee (2022).

reverses sign in the *Extroversion* and *Introversion* treatments as strong evidence that this correlation is influenced by the treatment—the availability of information. By contrast, IQ is uncorrelated with the incentivized extroversion score in the *No Priming* treatment.

For perspective on the magnitudes of the correlations between IQ and incentivized extroversion, consider subject pool trimming experiments similar to those discussed above. In the *Extroversion* (*Introversion*) treatment, 49 of 84 subjects (47 of 80 subjects) in the bottom half of the IQ distribution would be eliminated if half of the subjects were eliminated based on incentivized extroversion scores. That is, eliminating subjects based on incentivized extroversion scores does nearly as well eliminating low IQ subjects as it does eliminating subjects with undesirable non-incentivized extroversion scores. By contrast, in the *No Priming* treatment only 39 of 74 low IQ subjects would be eliminated based on incentivized extroversion scores—very nearly what one would expect if elimination were random.

For the personality traits that were not primed, we fail to reject the null hypothesis that the IQ coefficients are equal in the *Extroversion*, *Introversion*, and *No Priming* treatments when all of the remaining incentivized Big Five traits except openness serve as the dependent variable; for incentivized openness we reject this null at the 5% level.[35] IQ and incentivized openness are positively correlated in the *Extroversion* treatment, negatively correlated in the *Introversion* treatment, and essentially uncorrelated in the *No Priming* treatment. Overall, the absence of correlations between IQ and the non-primed trait measures other than openness is consistent with

[35] One possibility is that the treatments inadvertently indicate that openness is desired in the *Extroversion* treatment but not desired in the *Introversion* treatment. Comparing the words in the ads to trait-descriptive adjective lists (e.g., Goldberg, 1981), however, indicated that the *Introversion* ad has as many words associated with openness as the *Extroversion* ad. Alternatively, subjects may bring beliefs to the lab about the correlations between desired traits. Specifically, subjects may believe that firms looking for extroversion (introversion) are also looking for openness (non-openness). In this case, more intelligent subjects who identify extroversion cues would also manipulate openness responses.

the hypothesis that IQ influences incentivized personality measures through an information processing channel.

We also fail to reject the equality of the coefficients for most of the other characteristics in the *Extroversion*, *Introversion* and *No Priming* treatments using any of the incentivized personality measures as the dependent variable. We reject the equality of the self-monitoring coefficients across treatments in the incentivized extroversion regressions, but at the same time we also fail to reject that these coefficients are jointly equal to zero.

The locus of control coefficient estimates, however, appear to vary with the treatments. We reject the equality of the locus of control coefficients across treatments at the 10% level in the incentivized conscientiousness, openness and neuroticism regressions, which suggests that the information available to test-takers influences how locus of control beliefs affect their responses to incentivized personality measures. McGee and McGee (2016) found evidence consistent with the hypothesis that locus of control beliefs influence effort when the return to effort is uncertain but not when this return is known in a search experiment. A similar dynamic may be at work here. Locus of control beliefs are correlated with incentivized personality measures in the *No Priming* and *Introversion* treatments where priming is absent or subtle and counterintuitive and thus the return to faking effort uncertain but never correlated with incentivized measures in the *Extroversion* treatment where the return to faking is perhaps clearer to subjects.[36] That locus of control beliefs can be correlated with incentivized personality measures such as those used in hiring suggests another channel through which these beliefs may influence job search beyond those documented in Caliendo et al. (2015) and McGee (2015).[37]

[36] Whereas the job ad in the *Extroversion* treatment had 16 bullet points describing the ideal candidate, the ads in the *Introversion* and *No Priming* treatment had only 9 and 4 such bullet points, respectively.

[37] Given that the priming in the ads only involves extroversion, one could pool the data from the treatments for the traits other than extroversion when regressing the incentivized scores on non-incentivized scores and other

6. Conclusion

Using a within-subject laboratory experiment with incentives but not directives for misrepresentation, we find that incentivized and non-incentivized personality measures are only weakly to moderately correlated. Providing test-takers with information about desired personality traits has little or no effect on these correlations but does lead to correlations between IQ and incentivized measures of the trait being sought. Similarly, locus of control beliefs are correlated with incentivized conscientiousness and neuroticism measures in the *No Priming* and *Introversion* treatments when the return to faking may be unclear but are uncorrelated with incentivized measures in the *Extroversion* treatment when the incentives for faking are clearer. Otherwise we find little evidence that the measured characteristics influence faking.

A number of questions follow from the study. First, would our findings generalize to real hiring settings? Online personality testing has emerged as an important selection tool for employers, but evidence about what employers learn from such tests from within-subject studies is virtually non-existent. Using a lab experiment to investigate incentivized personality tests has obvious shortcomings: most subjects understood that they were not applying for real jobs, the incentives provided to subjects are not the same as those of actual job applicants, and our subject population may not resemble the applicant population for any given job. We attempted to minimize these limitations by using ads for a job to which young university graduates might apply, but we concede that they may limit the external validity of our findings. Laboratory

characteristics to benefit from greater statistical power. Doing so, we find only that locus of control beliefs are significant at the 5% level in the conscientiousness and neuroticism regressions, while risk aversion and impression management are significant at the 10% level in the agreeableness regression. Pooling the data, however, is problematic for at least two reasons. First, it obscures the potentially moderating effects of the information about desired personalities in the treatments on the relationships between personality traits and non-personality characteristics. Second, subjects may associate the primed trait (extroversion) with a bundle of other traits, in which case the priming pertains to all of the Big Five traits.

experiments, however, have an indispensable role to play studying faking and its correlates insofar as obtaining incentivized and non-incentivized personality measures from individuals in an actual hiring setting has proven exceptionally difficult.

Second, how correlated with worker performance are incentivized personality measures, and do these correlations depend on the job or the information provided to test-takers? Many studies of the criterion validity of personality tests where job performance is concerned use non-incentivized personality measures (e.g., Barrick and Mount, 1991; Hurtz and Donovan, 2000), while the evidence on the effectiveness of job testing in the economics literature comes from single firms hiring for well-defined positions (Autor and Scarborough, 2008; Hoffman et al., 2017). Our findings suggest that the correlations between incentivized personality measures and productivity could be influenced by correlations between these measures and IQ and locus of control beliefs. Field experiments varying the information available to job applicants and tracking employment outcomes for applicants to different types of jobs could make useful contributions in this regard.

Finally, do our findings extend to other environments in which personalities are assessed with incentives for misrepresentation such as cover letters, job interviews, LinkedIn profiles or even dating? Can individuals misrepresent their personalities in these venues? Do more intelligent individuals perform better because they are better at inferring what interviewers want to hear? Is faking even a problem in these settings? Faking may be easier in some settings than others. For instance, misrepresenting oneself in person over an extended period of interviewing may be much harder than faking on an online personality test. We leave these questions for future research.

References

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond. 2019. "Preferences for truth-telling." *Econometrica*, 87(4),1115-1153.

Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz. 2011. "Personality Psychology and Economics." *Handbook of the Economics of Education*, 4, 1-181.

Autor, David H., and David Scarborough. 2008. "Does job testing harm minority workers? Evidence from retail establishments." *The Quarterly Journal of Economics*, 123 (1), 219-277.

Barrick, Murray R., and Michael K. Mount. 1991. "The big five personality dimensions and job performance: a meta-analysis." *Personnel Psychology*, 44 (1), 1-26.

Becker, Thomas and Alan Colquitt. 1992. "Potential versus actual faking of a biodata form: analysis along several dimensions of item type." *Personnel Psychology*, 45, 389–406.

Birkeland, Scott A., Todd M. Manson, Jennifer L. Kisamore, Michael T. Brannick, and Mark A. Smith. 2006. "A meta-analytic investigation of job applicant faking on personality measures." *International Journal of Selection and Assessment*, 14 (4), 317-335.

Borghans, Lex, Huub Meijers, and Bas ter Weel, B. 2008. "The Role of Noncognitive Skills in Explaining Cognitive Test Scores." *Economic Inquiry*, 46(1), 2–12.

Boyce, Anthony S. 2005. "An investigation of faking: Its antecedents and impacts in applicant settings." *Unpublished MA thesis, Michigan State University, Lansing*.

Brenčič, Vera and Andrew McGee. 2023a. "Employers' demand for personality traits." *IZA Discussion Paper 16083*.

Brenčič, Vera and Andrew McGee. 2023b. "Demand for personality traits, tasks, and sorting." *Research in Labor Economics*, 51, forthcoming.

Caliendo, Marco, Deborah Cobb-Clark, and Arne Uhlendorff. 2015. "Locus of control and job search strategies." *Review of Economics and Statistics*, 97(1), 88-103.

Chen, Yuanyuan, Shuaizhang Feng, James J. Heckman, and Tim Kautz. 2020. "Sensitivity of self-reported noncognitive skills to survey administration conditions." *Proceedings of the National Academy of Sciences* (*PNAS*), 117(2), 931-935.

Cobb-Clark, Deborah A. and Stefanie Schurer. 2012. "The stability of big-five personality traits." *Economics Letters*, 115(1),11-15.

Dal Bo, Pedro, Frederico Finan, and Martin Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics*, 128(3), 1169-1218.

Davidson, Russell and James MacKinnon. 1993. *Estimation and Inference in Econometrics.* New York: Oxford University Press.

DeYoung, Colin G., Lena C. Quilty, and Jordan B. Peterson. 2007. "Between facets and domains: 10 aspects of the Big Five." *Journal of Personality and Social Psychology*, 93(5), 880.

Donovan, John, Stephen Dwight, and Gregory Hurtz. 2003. "An assessment of the prevalence, severity and verifiability of entry-level applicant faking using the randomized response technique." *Human Performance*, 16, 81–106.

Doyle, Douglas, C. Clifton Layne, Raymond Mendel, and David Shiek. 1977. "Effects of Manipulation on Examination of Grades on Locus of Control Beliefs." *Psychological Reports*, 41, 379-382.

Duckworth, A.L., Quinn, P.D., Lynam, D.R., Loeber, R. and Stouthamer-Loeber, M. 2011. "Role of test motivation in intelligence testing." *Proceedings of the National Academy of Sciences*, *108*(19), 7716-7720.

Dunn, Wendy, Michael Mount, Murray Barrick, and Deniz Ones. 1995. "Relative importance of personality and general mental ability in managers' judgments of applicant qualifications." *Journal of Applied Psychology*, 80 (4), 500–509.

Dwight, Stephen A. and John J. Donovan. 2003. "Do Warnings Not to Fake Reduce Faking?" *Human Performance*, 16(1), 1-23.

Ellingson, Jill, Paul Sackett, and Brian Connelly. 2007. "Personality Assessment Across Selection and Development Contexts: Insights Into Response Distortion." *Journal of Applied Psychology*, 92(2), 386-395.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental economics*, 10(2), 171-178.

Fischbacher, Urs and Franziska Föllmi-Heusi. 2013. "Lies in Disguise—An Experimental Study on Cheating." *Journal of the European Economic Association*, 11(3), 525–547.

Furnham, Adrian. 1990 "Faking personality questionnaires: Fabricating different profiles for different purposes." *Current Psychology*, 9(1), 46-55.

Furnham, Adrian. 1997. "Knowing and Faking One's Five-Factor Personality Score." *Journal of Personality Assessment*, 69(1), 229-243.

Ghiselli, Edwin E., and Richard P. Barthol. 1953. "The validity of personality inventories in the selection of employees." *Journal of Applied Psychology*, 37 (1), 18.

Gielen, Anne C. 2013. "Repeated job quits: stepping stones or learning about quality?" *IZA Journal of European Labor Studies*, (**2)**, 7.

Goldberg, Lewis R. 1981. "Language and individual differences: The search for universals in personality lexicons." *Review of Personality and Social Psychology*, 2 (1), 141-165.

Griffith, Richard, Tom Chmielowski, and Yukiko Yoshita. 2007. Do applicants fake? An examination of the frequency of applicant faking behavior." *Personnel Review*, 36(3), 341-355.

Heckman, James J., and Tim Kautz. 2012. "Hard evidence on soft skills." *Labour Economics*, 19 (4), 451-464.

Hoffman, Mitchell, Lisa B. Kahn, and Danielle Li. 2018. "Discretion in hiring." *Quarterly Journal of Economics*, 133(2), 765-800.

Holden, Ronald and Nicole Hibbs. 1995. "Incremental validity of response latencies for detecting fakers on a personality test." *Journal of Research in Personality*, 29, 362–372.

Holt, Charles A. and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review*, 92(5), 1644-1655.

Hough, Leaetta, Newell Eaton, Marvin Dunnette, John Kamp, and Rodney McCloy. 1990. "Criterion-related validities of personality constructs and the effect of response distortion on those validities." *Journal of Applied Psychology*, 75(5), 581-595.

Hough, Leaetta M., and Frederick L. Oswald. 2000. "Personnel selection: Looking toward the future--Remembering the past." *Annual Review of Psychology*, 51(1), 631-664.

Hough, Leaetta M., and Frederick L. Oswald. 2008. "Personality testing and industrial–organizational psychology: Reflections, progress, and prospects." *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*(3), 272–290.

Hu, Jing and Brian S. **Connelly.** 2021. "Faking by actual applicants on personality tests: A meta-analysis of within-subject studies." *International Journal of Selection and Assessment*, 29, 412-426.

Huber, Christopher R., Nathan R. Kuncel, Katie B. Huber, and Anthony S. Boyce. 2021. "Faking and the Validity of Personality Tests: An Experimental Investigation Using Modern Forced Choice Measures." *Personnel Assessment and Decisions*, 7 (1), 20-30.

Hurtz, Gregory M., and John J. Donovan. 2000. "Personality and job performance: The Big Five revisited." *Journal of Applied Psychology*, 85 (6), 869-879.

Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi. 1997. "The effects of human resource management practices on productivity: A study of steel finishing lines." *The American Economic Review*, 87 (3) , 291-313.

Isaacson, Joshua, Mitchell Peterson, Stephanie Turner, Richard Griffith, and Mei-Chuan Kung. 2009. "Investigating Response Distortion in Selection Settings: Evaluating Subgroup Differences." *24th Annual Conference of the Society for Industrial and Organizational Psychology.*

Jackson, Douglas N. 1994. *Jackson Personality Inventory-Revised manual*. Port Huron, MI: Sigma Assessment Systems.

Kleinmann, Martin, Pia V. Ingold, Filip Lievens, Anne Jansen, Klaus G. Melchers and Cornelius J. König. 2011. "A different look at why selection procedures work : The role of candidates' ability to identify criteria." *Organizational Psychology Review*, 1(2), 128–146.

König, Cornelius J., **Ann-Sibill Merz, and Nicoletta Trauffer**. 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.

Krammer, Georg. 2020. "Applicant Faking of Personality Inventories in College Admission: Applicants' Shift From Honest Responses is Unsystematic and Related to the Perceived Relevance for the Profession." *Journal of Personality Assessment*, 102 (6), 758-769.

Kroger, Rolf and Linda Wood. 1993. "Reification, "faking," and the Big Five." *American Psychologist*, 48, 1297-1298.

McGee, Andrew. 2015. "How the Perception of Control Influences Unemployed Job Search." *Industrial and Labor Relations Review*, 68(1), 184-211.

McGee, Andrew. ECIN Replication Package for "Whoever you want me to be: personality and incentives." Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2023-11-08. https://doi.org/10.3886/E194887V1

McGee, Andrew and Peter McGee. 2016. "Search, effort, and locus of control." *Journal of Economic Behavior & Organization*, 126, 89-101.

McGee, Andrew and Peter McGee. 2020. "Whoever you want me to be: personality and incentives." *IZA Discussion Paper 13809*.

McGee, Andrew and Peter McGee. 2022. "Gender and race differences on incentivized personality measures." *Working paper*.

McLeod, Beverly A., and Randy L. Genereux. 2008. "Predicting the acceptability and likelihood of lying: The interaction of personality with type of lie." *Personality and Individual Differences*, 45 (7), 591-596.

McFarland, Lynn A., and Ann Marie Ryan. 2000. "Variance in faking across noncognitive measures." *Journal of Applied Psychology*, 85 (5), 812.

Morgeson, Frederick, Michael Campion, Robert Dipboye, John Hollenbeck, Kevin Murphy, and Neal Schmitt. 2007a. "Reconsidering the use of personality tests in personnel selection contexts." *Personnel Psychology*, 60 (3), 683-729

Morgeson, Frederick P., Michael A. Campion, Robert L. Dipboye, John R. Hollenbeck, Kevin Murphy, and Neal Schmitt. 2007b. "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.

Ones, Deniz and Chockalingam Viswesvaran. 1998. "The effects of social desirability and faking on personality and integrity assessment for personnel selection." *Human Performance*, 11, 245–269.

Ones, Deniz S., Stephan Dilchert, Chockalingam Viswesvaran, and Timothy A. Judge. 2007. "In support of personality assessment in organizational settings." *Personnel Psychology*, 60 (4), 995-1027.

Paulhus, Delroy L. 1984. "Two-component models of socially desirable responding." *Journal of Personality and Social Psychology*, 46 (3), 598-609.

Paulhus, Delroy L. 2002. "Socially desirable responding: The evolution of a construct." *The role of constructs in psychological and educational measurement* 4969.

Raven, John C. 1998. *Raven's progressive matrices*. Oxford: Oxford Psychologists Press.

Raghavan, Manish, Solon Barocas, Jon Kleinberg, and Karen Levy. 2020. "Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices." In *Conference on Fairness, Accountability, and Transparency (FAT* '20*), January 27–30, 2020, Barcelona, Spain. ACM, New York, NY, USA, 13 pages.

Rosse, Joseph G., Mary D. Stecher, Janice L. Miller, and Robert A. Levin. 1998. "The impact of response distortion on preemployment personality testing and hiring decisions." *Journal of Applied Psychology*, 83 (4), 634.

Rothstein, Mitchell G. and Goffin, Richard D. 2006. "The use of personality measures in personnel selection: What does current research support?" *Human Resource Management Review*, 16(2), 155-180.

Rotter, Julian B. 1966. "Generalized Expectancies for Internal Versus External Control of Reinforcement." *Psychological Monographs General and Applied*, 80 (1, Whole No. 609).

Scheier, Michael F., Charles S. Carver, and Michael W. Bridges. 1994. "Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test." *Journal of Personality and Social Psychology* 67 (6), 1063.

Schmidt, Frank L., and John Hunter. 1998. "The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings." *Psychological Bulletin*, 124, 262-274.

Segal, Carmit. 2012. "Working when no one is watching: Motivation, test scores, and economic success." *Management Science*, 58 (8), 1438-1457.

Snyder, Mark. 1974. "Self-monitoring of expressive behavior." *Journal of Personality and Social Psychology*, 30, 526-537.

Tett, Robert P., Kurt A. Freund, Neil D. Christiansen, Kevin E. Fox, and John Coaster. 2012. "Faking on self-report emotional intelligence and personality tests: Effects of faking opportunity, cognitive ability, and job type." *Personality and Individual Differences*, 52(2), 195-201.

Vasilopoulos, Nicholas L., Jeffrey M. Cucina, and Julia M. McElreath. 2005. "Do warnings of response verification moderate the relationship between personality and cognitive ability?" *Journal of Applied Psychology*, 90(2), 306-322.

Velicer, Wayne and Barbara Weiner. 1975. "Effects of sophistication and faking sets on the Eysenck Personality Inventory." *Psychological Reports*, 37, 71–73.

Viswesvaran, Chockalingam and Deniz Ones. 1999. "Meta-analyses of fakability estimates: implications for personality measurement." *Educational and Psychological Measurement*, 59 (2), 197–210.

Weber, Lauren, and Elzabeth Dwoskin. 2014. "Are workplace personality tests fair? Growing use of tests sparks scrutiny amid questions of effectiveness and workplace discrimination." *The Wall Street Journal*, September 29, 2014.

Wehner, Caroline, Andries de Grip, and Harald Pfeifer. 2022. "Do recruiters select workers with different personalities for different tasks? A discrete choice experiment." *Labour Economics*, 78, 102186.

Ziegler, Matthias. 2011. "Applicant Faking: A Look Into the Black Box." *The Industrial-Organizational Psychologist*, 49(1), 29-36.

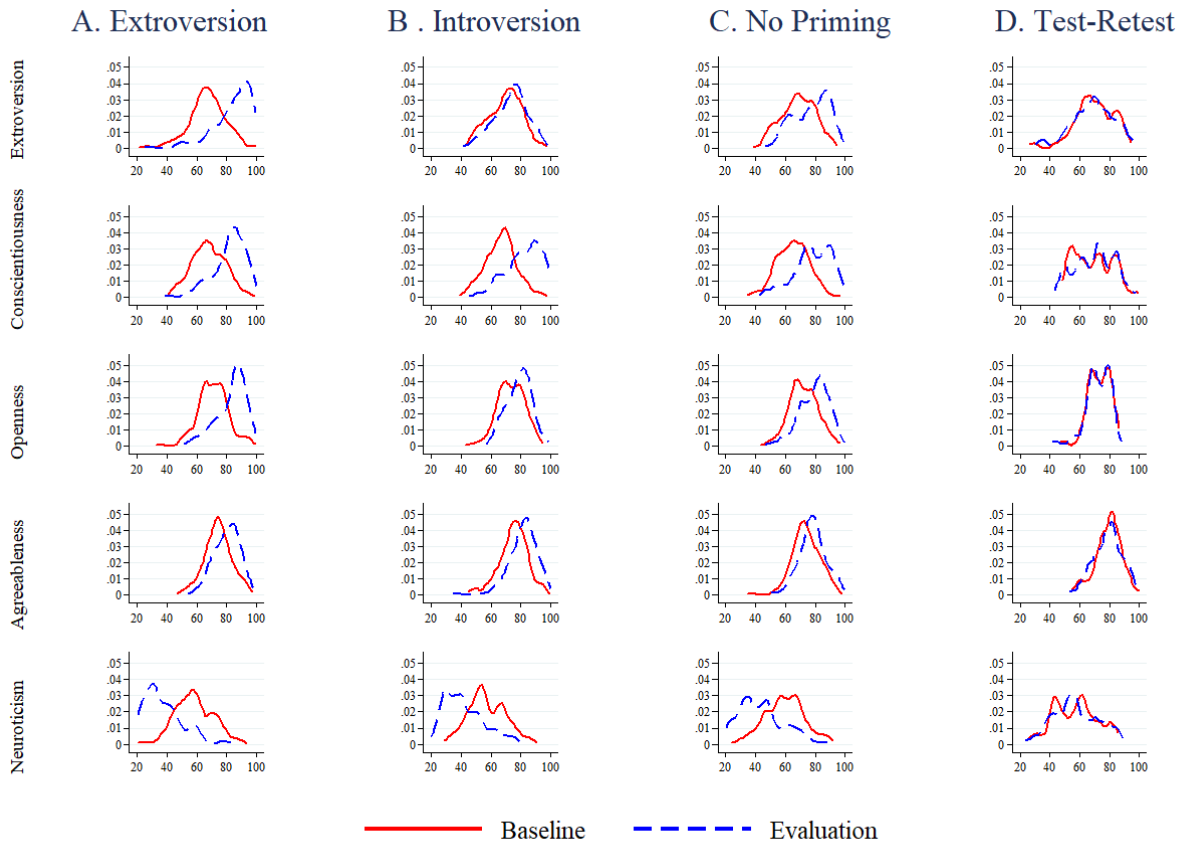

Figure 1: Kernel density plots

Notes: Kernel density plots of the raw scores from the Baseline and Evaluation sessions by treatment are displayed for the Big Five trait specified in each row. Column A depicts the densities for the *Extroversion* treatment, Column B the *Introversion* treatment, Column C the *No Priming* treatment, and Column D the *Test-Retest* sessions. The Evaluation session in the *Test-Retest* sessions was the same as the Baseline session and thus not incentivized. The p-values for the two-sample Kolmogorov-Smirnov tests for the equality of the extroversion distributions are 0.000, 0.029, and 0.000 for Columns A, B, and C, respectively. For all of the remaining traits, the p-values for the two-sample Kolmogorov-Smirnov tests are 0.000 in Columns A, B and C in every case. For the extroversion distributions, the estimated probability that a random draw from the Baseline distribution is larger than a random draw from the Evaluation distribution is 0.154 in Column A, 0.407 in Column B, and 0.300 in Column C. For the conscientiousness distributions, these probabilities are 0.175 in Column A, 0.167 in Column B, and 0.212 in Column C. For the openness distributions, these probabilities are 0.194 in Column A, 0.338 in Column B, and 0.290 in Column C. For the agreeableness distributions, these probabilities are 0.313 in Column A, 0.299 in Column B, and 0.361 in Column C. For the neuroticism distributions, these probabilities are 0.876 in Column A, 0.807 in Column B, and 0.784 in Column C.

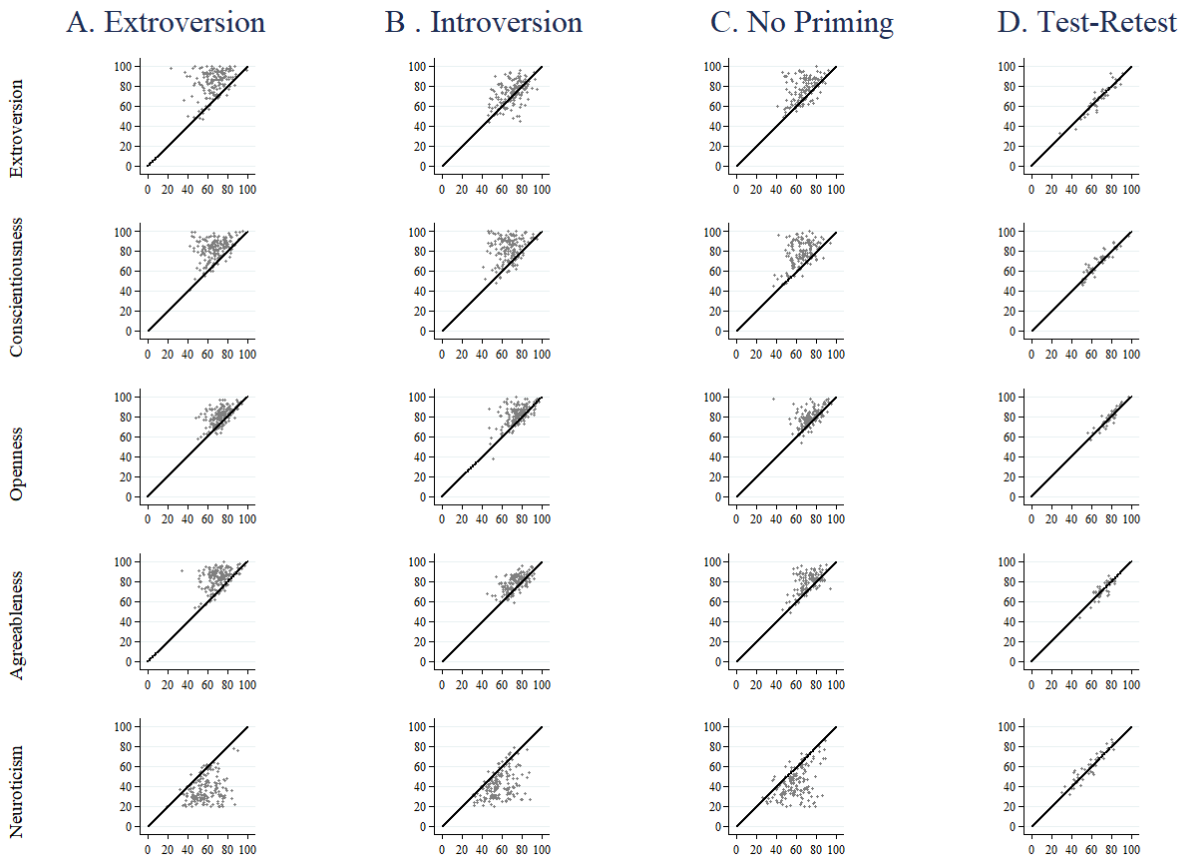

Figure 2: Scatterplots of Baseline against Evaluation Scores

Notes: The figures display scatterplots of the raw, Baseline scores on the x-axis against the raw, Evaluation scores on the y-axis by treatment for the Big Five trait specified in a given row. The plots in each panel also depict a 45-degree line as a reference. Column A depicts the scatterplots for the *Extroversion* treatment, Column B the *Introversion* treatment, Column C the *No Priming* treatment, and Column D the *Test-Retest* sessions. The Evaluation session in the *Test-Retest* sessions was the same as the Baseline session and thus not incentivized.

Table 1: Summary statistics for measures other than personality

	Treatment			
	Extroversion	Introversion	No Priming	KW p-value
	(1)	(2)	(3)	(4)
	A. Measured in the Baseline Session			
Male	0.44	0.49	0.51	0.46
Age	21.91	20.82	20.46	0.01
	(4.52)	(3.12)	(2.39)	
White	0.19	0.29	0.24	0.09
Self-deception	5.73	6.23	5.37	0.15
	(3.47)	(3.60)	(3.08)	
Impression management	6.34	5.82	5.86	0.50
	(3.83)	(3.46)	(3.38)	
Self-monitoring	27.34	28.88	28.34	0.27
	(8.33)	(8.67)	(7.96)	
Machiavellianism	18.32	19.34	18.84	0.17
	(4.65)	(4.64)	(4.61)	
Acceptability of lies	3.60	3.35	3.60	0.20
for personal gain	(1.61)	(1.44)	(1.35)	
Acceptability of lies to be kind	5.01	4.97	5.17	0.44
	(1.46)	(1.32)	(1.35)	
Acceptability of lies	4.78	4.59	4.79	0.28
to avoid conflict	(1.62)	(1.43)	(1.36)	
Acceptability of lies	4.13	3.84	4.20	0.09
to gain social acceptance	(1.69)	(1.52)	(1.53)	
Optimism	20.93	21.24	20.62	0.47
	(4.65)	(4.86)	(4.86)	
% potential unserious responders	0.08	0.05	0.07	0.48
	B. Measured in the Evaluation Session			
IQ	50.12	50.27	50.75	0.53
	(5.83)	(4.73)	(4.79)	
Risk aversion	5.64	5.41	5.33	0.50
	(1.98)	(2.02)	(2.12)	
Self-deception	7.08	6.51	6.44	0.65
	(4.72)	(4.19)	(3.83)	
Impression management	6.51	5.90	6.13	0.45
	(4.24)	(3.91)	(3.95)	
Optimism	22.61	22.26	21.93	0.44
	(5.00)	(4.81)	(4.61)	
Locus of control	11.65	11.63	11.72	0.93
	(2.15)	(1.94)	(2.01)	
# of personality tests taken for job	0.94	0.89	0.76	0.34
	(1.35)	(1.37)	(1.27)	
# of subjects	167	160	147	

Notes: Standard deviations in parentheses. Column (4) reports the p-values for the rank-based nonparametric Kruskal-Wallis test for significant differences across treatment for each variable.

Table 2: Personality trait scores with and without incentives

	Treatment			
	Extroversion	Introversion	No Priming	KW p-value
	(1)	(2)	(3)	(4)
			A. Extroversion	
Baseline	67.04	69.47	68.56	0.16
	(12.33)	(10.89)	(11.31)	
Evaluation	84.19	73.07	77.30	0.00
	(12.76)	(11.24)	(11.88)	
			B. Conscientiousness	
Baseline	67.76	66.42	65.03	0.11
	(11.11)	(10.32)	(10.71)	
Evaluation	82.05	81.96	78.06	0.01
	(10.79)	(11.84)	(12.53)	
			C. Agreeableness	
Baseline	74.12	74.86	73.80	0.31
	(9.12)	(9.69)	(9.26)	
Evaluation	80.06	81.36	78.27	0.00
	(8.46)	(9.35)	(8.86)	
			D. Openness/Intellect	
Baseline	70.95	73.34	72.03	0.05
	(9.98)	(8.96)	(9.52)	
Evaluation	82.98	78.51	79.21	0.00
	(9.99)	(8.13)	(9.79)	
			E. Neuroticism	
Baseline	58.35	57.26	59.36	0.26
	(12.77)	(12.42)	(13.16)	
Evaluation	37.27	41.73	43.71	0.00
	(12.72)	(13.36)	(14.76)	
# of subjects	167	160	147	

Notes: Each panel reports the mean non-incentivized (Baseline) and incentivized (Evaluation) scores for each personality trait by treatment. Columns (1) to (3) report the means for the *Extroversion*, *Introversion*, and *No priming* treatments, respectively. Column (4) reports the p-values for the rank-based nonparametric Kruskal-Wallis test for significant differences across treatments for the variable in the far-left column. The p-values for the nonparametric Wilcoxon signed-rank tests of equality of the means for the Baseline and Evaluation scores are less than 0.001 for each trait in every treatment. Standard deviations are given in parentheses.

Table 3: Correlations between trait scores with & without incentives

	Treatment			
	Extroversion	Introversion	No Priming	Retest 1-month/1-week
	(1)	(2)	(3)	(4)
		A. Pearson correlation coefficient		
Extroversion	0.38	0.45	0.50	0.83/0.94
Conscientiousness	0.35	0.21**	0.43	0.86/0.93
Agreeableness	0.55	0.53	0.49	0.79/0.92
Openness	0.42*	0.63	0.57	0.82/0.83
Neuroticism	0.30	0.45	0.44	0.85/0.94
		B. Spearman's rank correlation coefficient		
				1-week
Extroversion	0.31*	0.44	0.47	0.93
Conscientiousness	0.32	0.15*	0.33	0.94
Agreeableness	0.54	0.44	0.55	0.90
Openness	0.39	0.64**	0.48	0.77
Neuroticism	0.24	0.43	0.35	0.93
		C. Kendall's tau-b coefficient		
Extroversion	0.23	0.32	0.37	0.80
Conscientiousness	0.24	0.10	0.24	0.79
Agreeableness	0.40	0.33	0.43	0.78
Openness	0.28	0.46	0.36	0.60
Neuroticism	0.17	0.32	0.27	0.78
# of subjects	167	160	147	45

Notes: Columns 1 to 3 of Panel A report pairwise Pearson correlation coefficients for the non-incentivized and incentivized trait scores measured in the Baseline and Evaluation sessions, respectively, by treatment, while Panels B and C report the analogous Spearman rank correlation coefficients and Kendall's tau-b coefficients, respectively. In Panel A, Column 4 reports the retest correlations reported in DeYoung et al. (2007) at a 1-month interval and in our *Retest* sessions at a 1-week interval. In Panels B and C, Column 4 reports the statistics from the *Retest* sessions only. All of the correlations are statistically significant at the 1% level. Panels A, B and C also report the significance levels of z-tests of the equality of the correlations in the *Extroversion* (Column 1) and *Introversion* (Column 2) treatments with the correlation in the *No Priming* treatment (Column 3). Significance levels: *** p<0.01, ** p<0.05, * p<0.10

Table 4a: Incentivized extroversion regressions

	Treatment			p-values of
	Extroversion	Introversion	No Priming	(1)=···=(3)=0 [(1)=(2)=(3)]
	(1)	(2)	(3)	(4)
Non-incentivized extroversion	0.413***	0.483***	0.426***	0.000
	(0.135)	(0.115)	(0.114)	[0.909]
IQ	0.173***	-0.194***	0.048	0.001
	(0.057)	(0.074)	(0.087)	[0.000]
Risk aversion	0.158**	0.088	0.019	0.110
	(0.070)	(0.091)	(0.069)	[0.368]
Locus of control	-0.062	0.047	0.113	0.393
	(0.074)	(0.078)	(0.081)	[0.270]
Acceptability of lies for gain	0.030	0.048	-0.014	0.938
	(0.094)	(0.089)	(0.120)	[0.917]
Self-deception	0.075	0.034	-0.053	0.745
	(0.081)	(0.089)	(0.111)	[0.647]
Impression management	-0.139	0.093	0.023	0.424
	(0.099)	(0.106)	(0.116)	[0.256]
Self-monitoring	-0.109	-0.068	0.168*	0.167
	(0.089)	(0.117)	(0.093)	[0.080]
Machiavellianism	0.027	-0.008	-0.028	0.982
	(0.098)	(0.124)	(0.094)	[0.920]
Optimism	0.090	-0.032	0.000	0.780
	(0.092)	(0.088)	(0.093)	[0.616]
Constant	-0.001	0.015	-0.014	
	(0.077)	(0.075)	(0.076)	
R^2	0.237	0.264	0.285	

Notes: Columns (1) to (3) report coefficient estimates from separate OLS regressions for each treatment. The dependent variable in each regression is the incentivized extroversion score from the Evaluation session standardized using the mean and standard deviation of the incentivized scores from subjects in the same treatment. The non-incentivized extroversion score used in the controls is standardized using the mean and standard deviation for the non-incentivized scores from subjects in the same treatment. The remaining controls are standardized using the full sample of all subjects in all treatments. The heteroskedasticity-robust standard errors reported in parentheses are scaled by the square of the observation's variance estimate (i.e., the diagonal element of the projection matrix) as suggested by Davidson and MacKinnon (1993) given our sample sizes and implemented in Stata using the "vce(hc3)" option. Column (4) reports the p-values for heteroscedasticity-robust Wald tests for the hypotheses that the coefficients in Columns (1) to (3) are jointly significant and that the coefficients are equal when estimated in a fully-interacted model. Significance levels: *** p<0.01, ** p<0.05, * p<0.10

Table 4b: Incentivized conscientiousness regressions

	Treatment			p-value of
	Extroversion	Introversion	No Priming	(1)=···=(3)=0 [(1)=(2)=(3)]
	(1)	(2)	(3)	(4)
Non-incentivized conscientiousness	0.365***	0.183*	0.396***	0.000
	(0.113)	(0.109)	(0.104)	[0.320]
IQ	0.066	-0.045	0.004	0.765
	(0.068)	(0.099)	(0.083)	[0.630]
Risk aversion	0.104	-0.025	0.025	0.685
	(0.092)	(0.088)	(0.070)	[0.594]
Locus of control	-0.001	0.266***	0.179**	0.001
	(0.079)	(0.078)	(0.081)	[0.050]
Acceptability of lies for gain	0.038	0.066	-0.058	0.741
	(0.088)	(0.077)	(0.103)	[0.619]
Self-deception	0.020	0.017	0.012	0.994
	(0.100)	(0.107)	(0.094)	[0.998]
Impression management	-0.039	-0.037	-0.053	0.908
	(0.091)	(0.124)	(0.101)	[0.993]
Self-monitoring	0.025	-0.070	0.003	0.920
	(0.105)	(0.106)	(0.107)	[0.801]
Machiavellianism	0.054	0.189	0.011	0.455
	(0.099)	(0.124)	(0.109)	[0.540]
Optimism	-0.040	-0.122	0.010	0.562
	(0.102)	(0.089)	(0.092)	[0.583]
Constant	0.004	-0.001	-0.001	
	(0.081)	(0.082)	(0.078)	
R^2	0.152	0.139	0.223	

Notes: Columns (1) to (3) report coefficient estimates from separate OLS regressions for each treatment. The dependent variable in each regression is the incentivized conscientiousness score from the Evaluation session standardized using the mean and standard deviation of the incentivized scores from subjects in the same treatment. The non-incentivized conscientiousness score used in the controls is standardized using the mean and standard deviation for the non-incentivized scores from subjects in the same treatment. The remaining controls are standardized using the full sample of all subjects in all treatments. The standard errors reported in parentheses are scaled by the square of the observation's variance estimate (i.e., the diagonal element of the projection matrix) as suggested by Davidson and MacKinnon (1993) and implemented in Stata using the "vce(hc3)" option. Column (4) reports the p-values for heteroscedasticity-robust Wald tests for the hypotheses that the coefficients in Columns (1) to (3) are jointly significant and that the coefficients are equal when estimated in a fully-interacted model. Significance levels: *** p<0.01, ** p<0.05, * p<0.10

Table 4c: Incentivized agreeableness regressions

	Treatment			p-value of
	Extroversion	Introversion	No Priming	(1)=···=(3)=0 [(1)=(2)=(3)]
	(1)	(2)	(3)	(4)
Non-incentivized agreeableness	0.567***	0.714***	0.484***	0.000
	(0.078)	(0.112)	(0.161)	[0.422]
IQ	-0.032	0.075	0.136*	0.232
	(0.060)	(0.076)	(0.078)	[0.204]
Risk aversion	0.113	0.004	0.169**	0.041
	(0.079)	(0.063)	(0.067)	[0.190]
Locus of control	0.031	0.115	0.103	0.168
	(0.070)	(0.075)	(0.065)	[0.662]
Acceptability of lies for gain	0.016	0.111	0.058	0.577
	(0.065)	(0.088)	(0.101)	[0.684]
Self-deception	-0.015	-0.067	-0.025	0.896
	(0.073)	(0.095)	(0.110)	[0.908]
Impression management	-0.059	-0.194**	0.010	0.192
	(0.072)	(0.096)	(0.100)	[0.318]
Self-monitoring	-0.051	0.090	-0.065	0.680
	(0.080)	(0.106)	(0.105)	[0.495]
Machiavellianism	0.029	0.092	0.081	0.656
	(0.095)	(0.101)	(0.097)	[0.888]
Optimism	0.109	-0.084	0.003	0.338
	(0.079)	(0.069)	(0.073)	[0.186]
Constant	-0.009	0.001	-0.007	
	(0.070)	(0.068)	(0.079)	
R^2	0.326	0.375	0.294	

Notes: Columns (1) to (3) report coefficient estimates from separate OLS regressions for each treatment. The dependent variable in each regression is the incentivized agreeableness score from the Evaluation session standardized using the mean and standard deviation of the incentivized scores from subjects in the same treatment. The non-incentivized agreeableness score used in the controls is standardized using the mean and standard deviation for the non-incentivized scores from subjects in the same treatment. The remaining controls are standardized using the full sample of all subjects in all treatments. The standard errors reported in parentheses are scaled by the square of the observation's variance estimate (i.e., the diagonal element of the projection matrix) as suggested by Davidson and MacKinnon (1993) and implemented in Stata using the "vce(hc3)" option. Column (4) reports the p-values for heteroscedasticity-robust Wald tests for the hypotheses that the coefficients in Columns (1) to (3) are jointly significant and that the coefficients are equal when estimated in a fully-interacted model. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.10$

Table 4d: Incentivized openness regressions

	Treatment			p-value of
	Extroversion	Introversion	No Priming	(1)=···=(3)=0 [(1)=(2)=(3)]
	(1)	(2)	(3)	(4)
Non-incentivized openness	0.427***	0.602***	0.575***	0.000
	(0.105)	(0.080)	(0.100)	[0.397]
IQ	0.144**	-0.098	0.044	0.060
	(0.063)	(0.072)	(0.074)	[0.041]
Risk aversion	0.128	0.029	0.025	0.404
	(0.079)	(0.080)	(0.059)	[0.544]
Locus of control	-0.060	0.171**	0.044	0.135
	(0.069)	(0.081)	(0.074)	[0.096]
Acceptability of lies for gain	-0.010	0.061	-0.035	0.854
	(0.086)	(0.078)	(0.087)	[0.689]
Self-deception	-0.047	0.026	-0.071	0.838
	(0.092)	(0.084)	(0.102)	[0.728]
Impression management	-0.022	-0.013	-0.004	0.994
	(0.089)	(0.088)	(0.093)	[0.990]
Self-monitoring	0.120	-0.016	0.050	0.438
	(0.077)	(0.096)	(0.098)	[0.538]
Machiavellianism	-0.034	0.069	0.015	0.905
	(0.100)	(0.106)	(0.090)	[0.782]
Optimism	0.015	-0.126	-0.045	0.416
	(0.083)	(0.080)	(0.081)	[0.465]
Constant	0.006	0.006	-0.014	
	(0.074)	(0.064)	(0.073)	
R^2	0.233	0.442	0.334	

Notes: Columns (1) to (3) report coefficient estimates from separate OLS regressions for each treatment. The dependent variable in each regression is the incentivized openness score from the Evaluation session standardized using the mean and standard deviation of the incentivized scores from subjects in the same treatment. The non-incentivized openness score used in the controls is standardized using the mean and standard deviation for the non-incentivized scores from subjects in the same treatment. The remaining controls are standardized using the full sample of all subjects in all treatments. The standard errors reported in parentheses are scaled by the square of the observation's variance estimate (i.e., the diagonal element of the projection matrix) as suggested by Davidson and MacKinnon (1993) and implemented in Stata using the "vce(hc3)" option. Column (4) reports the p-values for heteroscedasticity-robust Wald tests for the hypotheses that the coefficients in Columns (1) to (3) are jointly significant and that the coefficients are equal when estimated in a fully-interacted model. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.10$

Table 4e: Incentivized neuroticism regressions

	Treatment			p-value of
	Extroversion	Introversion	No Priming	(1)=···=(3)=0 [(1)=(2)=(3)]
	(1)	(2)	(3)	(4)
Non-incentivized neuroticism	0.243**	0.426***	0.432***	0.000
	(0.107)	(0.084)	(0.111)	[0.341]
IQ	-0.127**	0.043	-0.051	0.201
	(0.064)	(0.077)	(0.082)	[0.234]
Risk aversion	-0.125	-0.071	-0.045	0.387
	(0.089)	(0.087)	(0.070)	[0.783]
Locus of control	0.017	-0.242***	-0.132	0.006
	(0.077)	(0.076)	(0.084)	[0.057]
Acceptability of lies for gain	-0.079	-0.120	0.039	0.381
	(0.086)	(0.083)	(0.113)	[0.521]
Self-deception	-0.025	-0.009	0.013	0.991
	(0.085)	(0.091)	(0.109)	[0.963]
Impression management	-0.058	0.126	0.013	0.569
	(0.088)	(0.100)	(0.105)	[0.389]
Self-monitoring	-0.030	0.109	0.014	0.726
	(0.094)	(0.100)	(0.111)	[0.590]
Machiavellianism	0.015	-0.198*	-0.085	0.292
	(0.097)	(0.113)	(0.106)	[0.362]
Optimism	-0.086	0.072	0.049	0.623
	(0.096)	(0.083)	(0.102)	[0.433]
Constant	0.013	-0.003	0.008	
	(0.081)	(0.073)	(0.081)	
R^2	0.137	0.302	0.220	

Notes: Columns (1) to (3) report coefficient estimates from separate OLS regressions for each treatment. The dependent variable in each regression is the incentivized neuroticism score from the Evaluation session standardized using the mean and standard deviation of the incentivized scores from subjects in the same treatment. The non-incentivized neuroticism score used in the controls is standardized using the mean and standard deviation for the non-incentivized scores from subjects in the same treatment. The remaining controls are standardized using the full sample of all subjects in all treatments. The standard errors reported in parentheses are scaled by the square of the observation's variance estimate (i.e., the diagonal element of the projection matrix) as suggested by Davidson and MacKinnon (1993) and implemented in Stata using the "vce(hc3)" option. Column (4) reports the p-values for heteroscedasticity-robust Wald tests for the hypotheses that the coefficients in Columns (1) to (3) are jointly significant and that the coefficients are equal when estimated in a fully-interacted model. Significance levels: *** p<0.01, ** p<0.05, * p<0.10

Appendix Table 1: Comparing returnees to non-returnees

Variable	Non-returnees (1)	Returnees (2)	p-value (3)
Male	0.49	0.48	0.88
Age	20.75	21.04	0.63
	(2.51)	(3.55)	
White	0.21	0.24	0.46
Self-deception	5.33	5.79	0.13
	(3.27)	(3.41)	
Impression management	5.70	6.01	0.46
	(3.30)	(3.57)	
Self-monitoring	29.14	28.17	0.22
	(7.88)	(8.34)	
Machiavellianism	18.97	18.83	0.68
	(4.43)	(4.64)	
Acceptability of lies for personal gain	3.53	3.52	0.92
	(1.34)	(1.48)	
Acceptability of lies to be kind	5.03	5.05	0.71
	(1.35)	(1.38)	
Acceptability of lies to avoid conflict	4.80	4.72	0.87
	(1.43)	(1.48)	
Acceptability of lies to gain social acceptance	4.17	4.05	0.47
	(1.61)	(1.59)	
Optimism	21.30	20.94	0.39
	(4.51)	(4.78)	
Extroversion	69.94	68.33	0.22
	(10.19)	(11.56)	
Conscientiousness	65.67	66.46	0.47
	(11.31)	(10.76)	
Agreeableness	75.01	74.27	0.53
	(9.79)	(9.35)	
Openness	72.74	72.09	0.41
	(9.43)	(9.53)	
Neuroticism	59.87	58.30	0.19
	(13.21)	(12.78)	
% potential unserious responders	0.09	0.07	0.42
# of subjects	162	474	

Notes: Standard deviations in parentheses. Non-returnees completed only a Baseline session. Returnees completed both a Baseline and Evaluation session in the *Extroversion*, *Introversion* and *No Priming* treatments. Column (3) reports the p-values for the nonparametric rank-sum tests of the hypothesis that returnee and non-returnee samples are from populations with the same distribution.

Appendix: Instructions, Job Descriptions, and Scales

Instructions (Baseline)

Please answer the following questions about yourself. Pay careful attention to response scales provided at the top of the screen as they may change from question to question.

You will be paid **$7** for participating in this experiment. Are there any questions?

Your participation in today's experiment makes you eligible for another experiment to be held in the future. You will receive an invitation to this experiment.

Instructions (Evaluation)

This is an experiment in the economics of decision making.

Each subject will be randomly assigned to a group of 3-5 subjects (most groups will have 4 subjects) ; each group will be assigned a job description. All members of a group will receive the same job description. We will then administer an intelligence test and a personality test to determine who to "hire" for the job. We will weigh these two tests according to some criteria and one subject within each group who best meets these criteria will receive a bonus of $25; the remaining subjects in each group will not receive any bonus.

After the bonus has been awarded, you will be asked to answer some further questions. You will be paid for one portion of the additional questions, which will be explained to you at the time.

Your earnings today will have four components.

1) You will be paid $0.20 for every correct answer on the intelligence test.
2) The subject who is "hired" from his or her group will receive a bonus payment of $25.
3) You will be paid for one portion of the additional questions.
4) All subjects who participate will receive a $7 show-up fee.

Are there any questions?

We give you a couple minutes now to read the job description before we begin the intelligence test. You will then have 20 minutes to complete the intelligence test.

Job descriptions (Extroversion)

Client Services Representative

About the Job

EXCELLENT OPPORTUNITY for a client services representative at a collegial and exciting staffing firm. This is an ideal position for an outgoing, friendly, energetic person who can represent our company to the outside world and help us stand out from the crowd.

Primary Duties:
Accomplish marketing and sales objectives by taking charge of programs to expand our customer base
Build relationships with key accounts by making regular visits
Maintain personal networks, participate in networking and professional organizations.
Implementation of business-need assessment programs for clients
Provide staffing support to clients to help them meet their business needs
Allocate firm resources across a diverse set of clients

An ideal candidate will be someone who:
Is a happy, cheerful, optimistic, and enthusiastic "go-getter"
Is proficient at building and maintaining relationships with client firms and staffers
Enjoys interacting with customers and the public
Is a fun-to-be-around person who customers will want to invite to backyard barbecues.
Thrives in crowds.
Has boundless energy and vigor to enthusiastically promote our firm at every opportunity.
Is aggressive and assertive in ensuring that our clients' problems get resolved in a timely fashion
Speaks up and takes charge to resolve problems
Is happy about working at our firm!

Is able to take initiative, be assertive, and follow projects to completion
Has great people skills
Is proficient in Microsoft Office (Word, Excel, PowerPoint, etc.)
Has excellent project management skills
Has strong analytical skills
Has strong communication skills
Has financial acumen

Skills/Qualifications:
Bachelor's Degree
Two years of experience in a related field
Valid driver's license upon starting work

Job descriptions (Introversion)

Client Services Representative

About the Job

EXCELLENT OPPORTUNITY for a client services representative at a low-key staffing firm. This is an ideal position for a contemplative, reserved, independent person who can represent our company and grow.

Primary Duties:

Accomplish marketing and sales objectives by developing programs that our customers implement remotely
Assist key accounts while working from home, often at off-peak times
Keeping up-to-date on industry developments without regular contact from supervisors or co-workers
Implementation of business-need assessment programs for clients
Provide staffing support to clients to help them meet their business needs
Allocate firm resources across a diverse set of clients

An ideal candidate will be someone who:
Enjoys tackling projects solo
Can get the job done without making a splash or interrupting clients normal business
Is capable of devising and revising business strategies independently
Gets a kick out of solving puzzles
Is proficient in Microsoft Office (Word, Excel, PowerPoint, etc.)
Has excellent project management skills
Has strong analytical skills
Has strong communication skills
Has financial acumen

Skills/Qualifications:
Bachelor's Degree
Two years of experience in a related field
Valid driver's license upon starting work

Job descriptions (No Priming)

Client Services Representative

About the Job
EXCELLENT OPPORTUNITY for a client services representative at a staffing firm.

Primary Duties:
Implementation of business-need assessment programs for clients
Provide staffing support to clients to help them meet their business needs

Allocate firm resources across a diverse set of clients

An ideal candidate will be someone who has:
Is proficient in Microsoft Office (Word, Excel, PowerPoint, etc.)
Has excellent project management skills
Has strong analytical skills
Has financial acumen

Skills/Qualifications:
Bachelor's Degree
Two years of experience in a related field
Valid driver's license upon starting work

Scales

Big Five

Neuroticism

+ keyed	Get angry easily. Get upset easily. Change my mood a lot. Am a person whose moods go up and down easily. Get easily agitated. Can be stirred up easily. Am filled with doubts about things. Feel threatened easily. Worry about things. Am easily discouraged. Become overwhelmed by events. Am afraid of many things.
– keyed	Seldom feel blue. Feel comfortable with myself. Rarely feel depressed. Am not embarrassed easily. Rarely get irritated. Keep my emotions under control. Rarely lose my composure. Am not easily annoyed.

Agreeableness

+ keyed	Feel others' emotions. Inquire about others' well-being. Sympathize with others' feelings. Take an interest in other people's lives. Like to do things for others. Respect authority. Hate to seem pushy. Avoid imposing my will on others. Rarely put people under pressure.
– keyed	Insult people. Believe that I am better than others. Take advantage of others. Seek conflict. Love a good fight. Am out for my own personal gain. Am not interested in other people's problems. Can't be bothered with other's needs. Am indifferent to the feelings of others. Take no time for others. Don't have a soft side.

Conscientiousness

+ keyed	Carry out my plans. Finish what I start. Get things done quickly. Always know what I am doing. Like order. Keep things tidy. Follow a schedule. Want everything to be "just right." See that rules are observed. Want every detail taken care of.

– keyed	Leave my belongings around. Am not bothered by messy people. Am not bothered by disorder. Dislike routine. Waste my time. Find it difficult to get down to work. Mess things up. Don't put my mind on the task at hand. Postpone decisions. Am easily distracted.

Extraversion

+ keyed	Make friends easily. Warm up quickly to others. Show my feelings when I'm happy. Have a lot of fun. Laugh a lot. Take charge. Have a strong personality. Know how to captivate people. See myself as a good leader. Can talk others into doing things. Am the first to act.
– keyed	Do not have an assertive personality. Lack the talent for influencing people. Wait for others to lead the way. Hold back my opinions. Am hard to get to know. Keep others at a distance. Reveal little about myself. Rarely get caught up in the excitement. Am not a very enthusiastic person.

Openness/Intellect

+ keyed	Am quick to understand things. Can handle a lot of information. Like to solve complex problems. Have a rich vocabulary. Think quickly. Formulate ideas clearly.
+ keyed	Enjoy the beauty of nature. Believe in the importance of art. Love to reflect on things. Get deeply immersed in music. See beauty in things that others might not notice. Need a creative outlet.
– keyed	Do not like poetry. Seldom get lost in thought. Seldom daydream. Seldom notice the emotional aspects of paintings and pictures. Have difficulty understanding abstract ideas. Avoid philosophical discussions. Avoid difficult reading material. Learn things slowly.

Scoring instructions:
For + keyed items, the response "Very Inaccurate" is assigned a value of 1, "Moderately Inaccurate" a value of 2, "Neither Inaccurate nor Accurate" a 3, "Moderately Accurate" a 4, and "Very Accurate" a value of 5.

For - keyed items, the response "Very Inaccurate" is assigned a value of 5, "Moderately Inaccurate" a value of 4, "Neither Inaccurate nor Accurate" a 3, "Moderately Accurate" a 2, and "Very Accurate" a value of 1.

Once numbers are assigned for all of the items in the scale, just sum all the values to obtain a total scale score.
DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*, 880-896

Machiavellianism

+ keyed	Find it easy to manipulate others. Have a natural talent for influencing people. Can talk others into doing things.
– keyed	Find it difficult to manipulate others. Hate being the center of attention. Lack the talent for influencing people.

Scoring instructions:
For + keyed items, the response "Very Inaccurate" is assigned a value of 1, "Moderately Inaccurate" a value of 2, "Neither Inaccurate nor Accurate" a 3, "Moderately Accurate" a 4, and "Very Accurate" a value of 5.

For - keyed items, the response "Very Inaccurate" is assigned a value of 5, "Moderately Inaccurate" a value of 4, "Neither Inaccurate nor Accurate" a 3, "Moderately Accurate" a 2, and "Very Accurate" a value of 1.

Once numbers are assigned for all of the items in the scale, just sum all the values to obtain a total scale score.
Jackson, D. N. (1994). *Jackson Personality Inventory-Revised manual*. Port Huron, MI: Sigma Assessment Systems

Self-monitoring

+ keyed		Would make a good actor.
		Put on a show to impress people.
		Am likely to show off if I get the chance.
		Am the life of the party.
		Am good at making impromptu speeches.
		Like to attract attention.
		Use flattery to get ahead.
– keyed		Hate being the center of attention.
		Would not be a good comedian.
		Don't like to draw attention to myself.

Scoring instructions:
For + keyed items, the response "Very Inaccurate" is assigned a value of 1, "Moderately Inaccurate" a value of 2, "Neither Inaccurate nor Accurate" a 3, "Moderately Accurate" a 4, and "Very Accurate" a value of 5.

For - keyed items, the response "Very Inaccurate" is assigned a value of 5, "Moderately Inaccurate" a value of 4, "Neither Inaccurate nor Accurate" a 3, "Moderately Accurate" a 2, and "Very Accurate" a value of 1.

Once numbers are assigned for all of the items in the scale, just sum all the values to obtain a total scale score.
Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, *30*, 526-537

Behavioral Inventory of Desirable Responding

My first impressions of people usually turn out to be right.
It would be hard for me to break any of my bad habits.
I don't care to know what other people really think of me.
I have not always been honest with myself.
I always know why I like things.
When my emotions are aroused, it biases my thinking.
Once I've made up my mind, other people can seldom change my opinion.
I am not a safe driver when I exceed the speed limit.
I am fully in control of my own fate.
It's hard for me to shut off a disturbing thought.
I never regret my decisions.
I sometimes lose out on things because I can't make up my mind soon enough.
The reason I vote is because my vote can make a difference.
My parents were not always fair when they punished me.
I am a completely rational person.
I rarely appreciate criticism.
I am very confident of my judgments
I have sometimes doubted my ability as a lover.
It's all right with me if some people happen to dislike me.
I don't always know the reasons why I do the things I do.
I sometimes tell lies if I have to.
I never cover up my mistakes.
There have been occasions when I have taken advantage of someone.
I never swear.
I sometimes try to get even rather than forgive and forget.
I always obey laws, even if I'm unlikely to get caught.
I have said something bad about a friend behind his/her back.
When I hear people talking privately, I avoid listening.
I have received too much change from a salesperson without telling him or her.
I always declare everything at customs.
When I was young I sometimes stole things.
I have never dropped litter on the street.
I sometimes drive faster than the speed limit.
I never read sexy books or magazines.
I have done things that I don't tell other people about.
I never take things that don't belong to me.
I have taken sick-leave from work or school even though I wasn't really sick.
I have never damaged a library book or store merchandise without reporting it.
I have some pretty awful habits.
I don't gossip about other people's business.

Self Deceptive Enhancement (SDE): Items 1 – 20 (Reverse scored items: 2,4,6,8,10,12,14,16,18,20).

Impression Management (IM): Items 21 – 40 (Reverse scored items: 21,23,25,27,29,31,33,35,37,39).
Scoring instructions:
Each + keyed item is scored on a 7-point Likert scale with "Not true" assigned a value of 1, "Somewhat true" assigned a value of 4, and "Very true" assigned a value of 7.

For - keyed items, the Likert ratings are reversed.

For each subscale, add one point for every 6 or 7, then sum the number of points
Paulhus, D. L. (1991). Balanced Inventory of Desirable Responding (BIDR) reference manual for version 6. *Manual available from author at Department of Psychology, University of British Columbia, Vancouver, BC, Canada V6T IY7.*

Abbreviated 4-item Rotter Internal-External Locus of Control Scale

A. What happens to me is my own doing.
B. Sometimes I feel that I don't have enough control over the direction my life is taking.

A. When I make plans, I am almost certain that I can make them work.
B. It is not always wise to plan too far ahead because many things turn out to be a matter of good or bad fortune.

A. In my case getting what I want has little or nothing to do with luck.
B. Many times we might just as well decide what to do by flipping a coin.

A. Many times I feel that I have little influence over the things that happen to me.
B. It is impossible for me to believe that chance or luck plays an important role in my life.

Scoring instructions:
Respondents choose which statement, either A or B, is closer to their opinion for each pair of statements. Respondents then choose whether the statement is "much closer" to their opinion or "slightly closer."

Choosing the external (fatalistic) statement and stating that it is "much closer" to their opinion is given a 1, choosing the external (fatalistic) statement and stating that it is "slightly closer" to their opinion is given a 2, choosing the internal (efficacious) statement and stating that it is "slightly closer" to their opinion is given a 3, and choosing the internal (efficacious) statement and stating that it is "much closer" to their opinion is given a 4. Total scores are the sum of the scores assigned to each pair of statements.

Optimism (Lot-R)

1. In uncertain times, I usually expect the best.

2. If something can go wrong for me, it will.

3. I'm always optimistic about my future.

4. I hardly ever expect things to go my way.

5. I rarely count on good things happening to me.

6. Overall, I expect more good things to happen to me than bad.

Scoring instructions:
The response "I DISagree a lot" is assigned a value of 1, "I DISagree a little" a value of 2, "I neither agree nor disagree" a 3, "I agree a little" a 4, and "I agree a lot" a value of 5.

Once numbers are assigned for all of the items in the scale, just sum all the values to obtain a total scale score.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, 67, 1063-1078

Holt-Laury (Low Stakes)

In the questions that follow, you are going to be asked to make ten decisions. Each decision will be between Option A and Option B. Please enter your decisions below and on the corresponding sheet that was handed out to you. Only one of the ten choices you make will be used to determine your earnings for this part of the experiment. After you answer all 10 questions you will be shown the "decision selected" and "outcome" which will be used to calculate your earnings. Be sure to write these down. Each decision is a paired choice between "Option A" and "Option B." You will make ten choices. Before you start making your ten choices, let me explain what these choices mean. Imagine a ten-sided die that will be used to determine payoffs; the faces are numbered from 1 to 10. After you have made all of your choices, the die would be thrown twice, once to select one of the ten decisions to be used, and a second time to determine what your payoff is for the option you chose, A or B, for the particular decision selected. Given this, you should make the choice that you would prefer if we were throwing the die for real. Now, please look at Decision 1 at the top. Option A pays 200 pennies if the throw of the ten sided die is 1, and it pays 160 pennies if the throw is 2-10. Option B yields 385 pennies if the throw of the die is 1, and it pays 10 pennies if the throw is 2-10. The other Decisions are similar, except that as you move down the table, the chances of the higher payoff for each option increase. In fact, for Decision 10 in the bottom row, the die will not be needed since each option pays the highest payoff for sure, so your choice here is between 200 pennies or 385 pennies.

To summarize, you will make ten choices: for each decision row you will have to choose between Option A and Option B. You may choose A for some decision rows and B for other rows, and you may change your decisions and make them in any order.

	Option A	Option B	Your Choice		
1.	1/10 of $2.00	9/10 of $1.60	1/10 of $3.85	9/10 of $0.10	A / B
2.	2/10 of $2.00	8/10 of $1.60	2/10 of $3.85	8/10 of $0.10	A / B
3.	3/10 of $2.00	7/10 of $1.60	3/10 of $3.85	7/10 of $0.10	A / B
4.	4/10 of $2.00	6/10 of $1.60	4/10 of $3.85	6/10 of $0.10	A / B
5.	5/10 of $2.00	5/10 of $1.60	5/10 of $3.85	5/10 of $0.10	A / B
6.	6/10 of $2.00	4/10 of $1.60	6/10 of $3.85	4/10 of $0.10	A / B
7.	7/10 of $2.00	3/10 of $1.60	7/10 of $3.85	3/10 of $0.10	A / B
8.	8/10 of $2.00	2/10 of $1.60	8/10 of $3.85	2/10 of $0.10	A / B
9.	9/10 of $2.00	1/10 of $1.60	9/10 of $3.85	1/10 of $0.10	A / B
10.	10/10 of $2.00	0/10 of $1.60	10/10 of $3.85	0/10 of $0.10	A / B

Lying scenarios

A co-worker of Melinda is hosting a party and asks Melinda if she is enjoying the food. In order not to hurt his feelings, Melinda lies and says the food is fantastic, even though it is overcooked and tasteless.

Mike is working on a group project with another student who comes up with an idea for their project. In order to avoid conflict, Mike lies and says he likes the idea, even though he thinks the idea is a poor one.

On a visit to another country, Lea buys some gold jewelry. In order to avoid paying duty on the jewelry, Lea lies and tells the customs official that she did not buy anything while in the country.

Sean accidentally backs into a parked car. As he is driving away, the owner arrives and asks Sean if he saw who damaged his car. In order to avoid paying for the damage, Sean lies and says he has no idea who did it.

Jamie's friend really wants her to go to a concert next weekend. Jamie would rather spend the weekend on her own. In order to avoid causing a conflict with her friend, Jamie lies and says she has to work that weekend.

Michelle's co-worker is very upset about a new policy at work. In order to avoid an argument, Michelle lies and agrees that the policy is unreasonable, even though she strongly approves of the policy.

Harry is overburdened at work and has little time for his family. A new co-worker asks Harry if he has time to help him learn the new bookkeeping system. In order to help him out, Harry lies and says he has time, even though he doesn't.

Mary's new co-worker asks for advice on applying for a position that has opened up in the company. Because Mary hopes to get the position herself, she lies and says the position has already been promised to someone else.

Kate's fellow students are complaining about an instructor they don't like. In order to fit in, Kate lies and says she dislikes the instructor as well, even though she really likes the instructor.

Tom and a friend buy a lottery ticket together. When Tom takes the ticket in to check if they won anything, he receives $200. In order to keep all the money for himself, Tom lies to his friend and tells him they won nothing.

Terry's new friend hates hunting. In order to be liked by her, Terry lies and tells her he has never hunted, even though he is an avid hunter.

Bob's neighbour asks if he will vote for him in the upcoming election. In order to avoid conflict, Bob lies and says he will, even though he intends to vote for another candidate.

One day Jerry is drinking with some co-workers who start talking about their experiences playing hockey. In order to fit in Jerry lies and tells stories about playing hockey himself, even though he has never actually played hockey.

Brad's friend asks if he will help her move the next day. In order to be helpful, Brad lies and tells her that he has nothing planned and will help her move, even though he had booked in to work that day.

Kira's friend tells her she really likes a new political party and asks Kira if she likes the party. In order to gain her friend's approval, she lies and says she does like them, even though she really dislikes the party.

Susan's friend at work asks Susan to write a letter of reference for her. To help her friend out, Susan lies in the letter and says she believes her friend is perfect for the job, even though Susan has some reservations about her friend's ability to do the job.

Scoring instructions:
Respondents rated how acceptable it was for the person in the scenario to have lied using nine-point Likert scales ranging from extremely unacceptable to extremely. For each respondent, one acceptability score for each type of lie (altruistic, conflict avoidance, social acceptance, and self-gain) was obtained by calculating his/her average score across the four relevant scenarios.

Scenarios 1,7, 14, and 16 are altruistic lies, scenarios 2, 5, 6, and 12 are lies to avoid conflict, scenarios 9, 11, 13, and 15 are lies to gain social acceptance, and scenarios 3, 4, 8, and 10 are lies told to benefit the liar.

McLeod, B. A., & Genereux, R. L. (2008). Predicting the acceptability and likelihood of lying: The interaction of personality with type of lie. *Personality and individual differences*, *45*(7), 591-596.

Department of Economics, University of Alberta
Working Paper Series

2023-11: Gender Differences in Reservation Wages in Search Experiments – **McGee, A.,** McGee, P.

2023-10: Designing Incentive Regulation in the Electricity Sector – **Brown, D.,** Sappington, D.

2023-09: Economic Evaluation under Ambiguity and Structural Uncertainties – **Andrews, B.**

2023-08: Show Me the Money! Incentives and Nudges to Shift Electric Vehicle Charge Timing – Bailey, M., **Brown, D.,** Shaffer, B., Wolak, F.

2023-07: Screening for Collusion in Wholesale Electricity Markets: A Review of the Literature – **Brown, D., Eckert, A., Silveira, D.**

2023-06: Information and Transparency: Using Machine Learning to Detect Communication – **Brown, D.,** Cajueiro, D., **Eckert, A., Silveira, D.**

2023-05: The Value of Electricity Reliability: Evidence from Battery Adoption – **Brown, D.,** Muehlenbachs, L.

2023-04: Expectation-Driven Boom-Bust Cycles – **Brianti, M.,** Cormun, V.

2023-03: Fat Tailed DSGE Models: A Survey and New Results – **Dave, C.,** Sorge, M.

2023-02: Evaluating the Impact of Divestitures on Competition: Evidence from Alberta's Wholesale Electricity Market – **Brown, D., Eckert, A.,** Shaffer, B.

2023-01: Gender Specific Distortions, Entrepreneurship and Misallocation – **Ranasinghe, A.**

2022-12: The Impact of Wholesale Price Caps on Forward Contracting – **Brown, D.,** Sappington, D.

2022-11: Strategic Interaction Between Wholesale and Ancillary Service Markets – **Brown, D., Eckert, A., Silveira, D.**

2022-10: Employing Gain-Sharing Regulation to Promote Forward Contracting in the Electricity Sector – **Brown, D.,** Sappington, D.

2022-09: Enhanced Intergenerational Occupational Mobility through Trade Expansion: Evidence from Vietnam – Mitra, D., Pham, H., **Ural Marchand, B.**

2022-08: Reputation of Quality in International Trade: Evidence from Consumer Product Recalls – **Zhong, J.**

2022-07: When Opportunity Knocks: China's Open Door Policy and Declining Educational Attainment – Jiang, X., Kennedy, K., **Zhong, J.**

2022-06: Asymmetric Impact of Real Effective Exchange Rate Changes on Domestic Output Revisited: Evidence from Egypt – **Sharaf, M.,** Shahen, A.

2022-05: Country-Based Investing with Exchange Rate and Reserve Currency – **Galvani, V.**

2022-04: The Mean-Variance Core of Cryptocurrencies: When More is Not Better – **Galvani, V.,** Faychuk, V.

2022-03: Outliers and Momentum in the Corporate Bond Market – **Galvani, V.,** Li, L.

2022-02: Higher Education Expansion and Crime: New Evidence from China – **Liu, X.,** Wang, C., Yan, Z., Zhao, Y.

2022-01: A Cautionary Tale of Fat Tails – **Dave, C.,** Dressler, S., Malik, S.

2021-14: Procurement Auctions for Regulated Retail Service Contracts in Restructured Electricity Markets – **Brown, D., Eckert, A.,** Olmstead, D.

2021-13: Socioeconomic and Demographic Disparities in Residential Battery Storage Adoption: Evidence from California – **Brown, D.**