This is the published version of a paper published in *Scandinavian Journal of Psychology*.

Citation for the original published paper (version of record):

Sinclair, S., Agerström, J. (2020)
Does expertise and thinking mode matter for accuracy in judgments of job applicants’ cognitive ability?
*Scandinavian Journal of Psychology*
https://doi.org/10.1111/sjop.12638

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-93044
Cognition and Neurosciences

Does expertise and thinking mode matter for accuracy in judgments of job applicants’ cognitive ability?

SAMANTHA SINCLAIR* and JENS AGERSTRÖM

Linnaeus University, Växjö, Sweden


The present research examined the role of thinking mode for accuracy in recruiters and laypeople’s judgments of applicants’ cognitive ability. In Study 1, students who relied on their intuition were somewhat less accurate. In Study 2, an experimental manipulation of thinking mode (intuitive vs analytical) revealed no apparent differences in accuracy. Moreover, there were no differences in accuracy or agreement between recruiters and laypeople. Examination of the use of specific resume content suggested that intuitive thinking corresponds to basing one’s judgments more on the way that applicants present themselves in their personal letter and less on diagnostic biographical information such as SAT scores. The findings point to the possibility that professional recruiters may not possess intuitive expertise in this context.

Key words: Accuracy, expertise, intuition, performance judgments, personnel selection.

*Samatha Sinclair, Department of Psychology, Linnaeus University, Universitéitsplatsen 1, SE-351 95, Växjö, Sweden. Tel: +46725941757; fax +0480-44 60 32. e-mail: samantha.sinclair@lnu.se

INTRODUCTION

Resume screening is one of the most frequently used selection methods (Dipboye & Jackson, 1999), most likely because it is a cheap and easy means of reducing the number of applicants who are subjected to more costly selection measures such as interviews. In this initial screening process, recruiters take on the role of an employment gatekeeper, determining which applicants are allowed to remain for further consideration and which ones become rejected from the organization (Cole, Feild, Giles & Harris, 2009). Previous research confirms that recruiters form impressions of applicants’ traits and skills based on reviewing their resumes, and that these impressions influence their judgments of applicant employability (Burns, Christiansen, Morris, Periard & Coaster, 2014; Cole et al., 2009). Inaccurate perceptions of applicants’ competence at this initial stage are therefore likely to result in suboptimal selection decisions, where the organization fails to hire the most competent applicant.

Given the important task placed upon recruiters, many organizations put considerable effort into finding recruiters with suitable training and experience on the job. It is typically assumed that such experience translates into valid selection judgments, but what evidence is there for this assumed link between experience and accuracy in judgments of applicants’ competence? Surprisingly, research on this topic is lacking, as past studies focus on accuracy in judgments of applicants’ personality. Considering that cognitive ability is one of the strongest predictors of work performance across many different occupations (Robertson & Smith, 2001), coupled with the central role it plays in candidate evaluation (Huffcutt, Conway, Roth & Stone, 2001), more research on the antecedents of accuracy in judgments of job applicant cognitive ability is called for.

EXPERTISE, THINKING MODE, AND ACCURACY

Most scholars now agree that information is processed in parallel, with one system responsible for conscious deliberating (System 2) and another for non-conscious intuition (System 1). The thinking mode of System 1 is associative, fast, affective, and holistic whereas slower, rule-based processes that demand more cognitive effort characterize the thinking mode of System 2 (Kahneman, 2003; Sloman, 1996). There is an ongoing debate about when one system has an advantage over the other. Whereas some studies find that the intuitive thinking mode produce impaired decisions (e.g., MacGregor, Lichtenstein & Slovic, 1988; Tversky & Kahneman, 1974), others find that it yields similar or better decisions compared to the analytical mode (e.g., Lee, Amir & Ariely, 2009). Hammond, Hamm, Grassia and Pearson (1987) suggested that System 1 is especially suited for tasks that include high familiarity, pictorial presentation, subjective measures, and unavoidability of an explicit organizing principle or algorithm to integrate cues, while analytical thinking is suitable for tasks that involve quantitative presentation, objective measures, and have an organizing principle readily available. Examples include exact mathematical tasks (Sinclair, 2010).

Several researchers on the topic of expertise (Benner & Tanner, 1987; Dreyfus & Dreyfus, 1986) have claimed that experts’ decisions are fast and intuitive because of extended experience with a task or in a domain. Some authors have suggested that intuiting boils down to a deliberation process that is inferential in nature and that occurs too fast to be consciously noticed (Klein, 1998; Pretz, 2011). This would imply that individuals who are not true experts would make poor decisions when relying on their intuition. However, a recent meta-analysis (Kuncel, Klieger, Connelly & Ones, 2013) showed that there is considerable loss of validity when data about applicants are subjectively and holistically combined, compared to when mechanical methods are...
used in employee selection — even by experts who are highly familiar with the jobs and organizations in question. Adding intuition to mechanical methods may not help either. For example, as early as 1943, a study by Sarbin found that academic achievement was better predicted by standardized scores alone than by counselors’ clinical judgments added to these scores. More recently, Kausel, Culbertson, and Madrid (2016) showed that recruiters who were presented with information from interviews suffered from more overconfidence than recruiters presented with test scores only. According to Kuncel et al. (2013), the most likely source of inferior accuracy of holistic/clinical methods is poorer reliability due less consistency, and a higher rate of unsystematic errors.

Interestingly, another common finding in studies on expertise is that longer experience on the job is not always associated with improved performance, for example, in clinical settings (Choudhry, Fletcher & Soumerai, 2005; Ericsson, 2007). The few studies that have compared recruiters and laypeople’s skill in judging applicants suggest that recruiters may not be superior to laypeople. One study (Schmid Mast, Bangerter, Bulliard & Aerni, 2011) found that while recruiters were better than students at detecting applicants’ global personality profile and deception, students outperformed recruiters on judging specific personality traits in several cases. Moreover, associations between both number of years of recruiters’ job experience and training and accuracy in personality and deception detection were inconsistent. A recent study of judgments of CEOs’ leadership ability from non-verbal facial cues (Re & Rule, 2016) found that professionals with no business experience even outperformed professionals with extensive business experience on this task.

There are several possible explanations for why recruiter experience does not lead to improved accuracy in selection judgments. For one thing, previous research attests that recruiters are by no means immune to bias in their perceptions and judgments of applicants (e.g., Dipboye, 1982; Morgeson & Campion, 1997; Posthuma, Morgeson & Campion, 2002). Reduced accuracy as a function of experience could presumably occur because being familiar with a domain may introduce an “illusion of knowledge” where idiosyncratic experiences are overgeneralized, in turn reducing reliance on more reliable features such as statistical cues. This would result in diminished accuracy while boosting confidence in experienced individuals’ judgments (Hall, Ariss & Todorov, 2007). Previous research suggests that confidence levels tend to increase rapidly when individuals become acquainted with a task (Sanchez & Dunning, 2018) and overconfidence can thus develop, meaning reduced accuracy, possibly through overreliance on intuition.

A key explanation may also be found in the type of learning environment. According to Kahneman and Klein’ (2009), certain criteria need to be present for intuitive expertise to develop. Specifically, the learning environment needs to be regular and predictable enough. It further needs to provide feedback that is both rapid and unequivocal. Indeed, studies of superior performance in domains such as music and athletics typically find that expertise has less to do with length of experience, and more to do with duration of deliberate practice in situations where immediate, accurate feedback is available, and thus opportunities for gradual improvement are present (Ericsson, Krampe & Tesch-Romer, 1993). For example, Crow (1957) found that medical students who received instructions, but no practice or feedback, were less accurate in their personality judgments than students who did not receive any instructions at all.

Although we know that applicants’ future job performance can be predicted (Millward, 2005), it is less clear whether recruiters typically receive rapid, accurate, and unambiguous feedback on their selection decisions. If the available evidence is often irregular and ambiguous, inaccurate judgments might be the rule rather than the exception (Swets, Dawes & Monahan, 2000). To our knowledge, only one study (Cole, Feild & Giles, 2003) has previously addressed the question of accuracy in recruiters’ judgments of applicant cognitive ability. However, that study did not include laypeople to which recruiters’ accuracy levels could be compared. Whether there are differences between recruiters and laypeople in this regard thus remains unknown. We put this question to the test in the present research. We also experimentally manipulate thinking mode to assess whether it has a causal effect on accuracy. Previous correlational research suggests that recruiters often tend to trust their “gut feelings” about applicants (Lodato, Highhouse & Brooks, 2011; Miles & Sadler-Smith, 2014) despite being taught to avoid this in their training (Woehr & Huffcut, 1994), but experimental research on the effects of thinking mode in this context is lacking.

In addition to examining the importance of expertise and thinking mode for accuracy, we focused on the role of specific resume content. Biographical information comprises historical and verifiable information, such as work experience and education, and this information is reported in resumes (Asher, 1972). Recruiters often assume that biographical information reveals behavioral patterns that convey important information about applicant skill and ability (e.g., Cole et al., 2009). Cole et al. (2003) had recruiters assess the extent to which specified items (e.g., volunteering for community service) were present in the resumes of job applicants, and then examined whether these judgments predicted the applicants’ cognitive ability. We extend this work by exploring whether self-reported reliance on valid cues in biographical information predicts accuracy in judgments of applicant cognitive ability. We were also interested in whether reliance on intuition relates to a focus on certain types of biographical information.

SUMMARY OF RESEARCH AIDS

Previous research on accuracy in recruiters’ judgments of applicants tends to focus on perceived personality. Although this work is important, it does not provide a complete picture of recruiters’ judgments. First, most studies have focused on either recruiters only (e.g., Apers & Derous, 2017; Cole et al., 2003, 2009) or on laypeople/students (e.g., Barrick, Patton & Haugland, 2010; Cole, Feild & Stafford, 2005), while research which compares these two groups to each other is scarce. Second, accuracy has typically been calculated by comparing recruiters’ ratings with applicants’ self-ratings of their personality, which can introduce problems with social desirability and self-presentation concerns (e.g., Hogan, 1991). Furthermore, cognitive ability tends to be a stronger predictor of work performance, compared to...
personality (Robertson & Smith, 2001), and is thus interesting in its own right.

The first aim of the present research was to examine if recruiters are more accurate than laypeople in predicting applicants’ performance on a cognitive test. Cognitive ability was described as highly important for the position that the applicant had applied for. We operationalized (in)accuracy as the deviation in participants’ predictions of applicant performance from number of questions that the applicant answered correctly in the cognitive ability test. We focus on relative rather than absolute accuracy; that is, we examine factors that may improve or reduce accuracy, rather than concluding whether judgments are accurate or inaccurate per se.

A second aim was to look at the role of intuitive versus analytical thinking, and its possible interaction with expertise, for accuracy levels in these judgments. Third, we sought to investigate whether reliance on intuition corresponds to a focus on certain types of resume content, and if more attention to valid cues corresponds to improved accuracy. Fourth, we looked at how certain types of resume content, and if more attention to valid cues corresponds to improved accuracy. Because previous research on this topic is scarce, we approach some of our research questions in an exploratory manner. For reasons of transparency, we report how we determined the sample size, all data exclusions (if any), all manipulations, and all measures in the current research. We planned the sample size of both studies to have at least 80% power to detect medium-sized effects.

We first conducted a pilot study to develop applications for a fictive job and to confirm that that is possible to predict applicant cognitive ability from the resume content. We proceeded by conducting a correlational study to examine if students’ thinking mode and attention to specific types of resume content predict accurate judgments of applicant cognitive ability. Finally, we conducted an experimental study on professional recruiters and laypeople to address our main research questions.

STUDY 1
Considering past research showing that forming impressions of applicants in a subjective manner typically results in less valid judgments (Kuncel et al., 2013), we hypothesized that students enrolled in business and psychology majors, who are likely to have some knowledge and interest in the subject but are low in experience, would benefit from using an analytic rather than intuitive approach. Further, we expected individuals who are more confident to be prone to rely on their intuition. We also tested the assumption that relying on one’s intuition translates into focusing less on resume content that is diagnostic of cognitive performance (SAT scores), and more on the holistic impression based on the applicant’s personal letter. We expected attention to more valid cues in the resumes to predict accuracy. Finally, we expected judgments of applicant performance (regardless of accuracy) to predict employability judgments.

Method
Pilot study. Because we were interested in using authentic applicants and in having access to their performance on a cognitive task, 16 students served as bogus applicants. They first wrote applications to a fictive position as research assistant in a large research project. We chose this position as it is relevant for students from various backgrounds, allowing us to sample students from several disciplines (see Waung, Hynes & Beatty, 2014, for a similar procedure).

The “job applicants” were provided with information about the project and told that their applications were to be used in a future study where recruiters would guess their performance on a cognitive test. We asked them to complete a resume (CV and personal letter) as if they were applying for the job in real life. The CV included information about age, gender, previous work experience, education, language and computer skills, grade point average, acquired math level, the Swedish equivalent to the SATs, leadership experience, and the possibility to list other qualifications that the applicant deemed relevant. This content is similar to Brown and Campion’s (1994) frequently used resume taxonomy. In the personal letter, we asked the applicants to introduce themselves and describe their background, personality, personal strengths, goals, and leisure activities.

Next, they rated their perceived skills in the areas covered by the cognitive test, such as creative problem solving and probability reasoning. They were then provided with brief information about the upcoming test and asked to estimate how many of the 11 questions in the test that they would succeed in answering correctly.

For our cognitive ability test, we used tasks that have been validated in previous research and we sought to cover a wide selection of cognitive abilities. Specifically, the applicants completed 11 cognitive tasks, of which two comprised creative problem solving tasks, including “A dealer in antique coins got an offer to buy a beautiful bronze coin. The coin had an emperor’s head on one side and the date 544 B.C. stamped on the other. The dealer examined the coin but instead of buying it, he called the police. Why?” (Schooher, Ohlsson & Brooks, 1993). The test also included one probability reasoning task (Kirkpatrick & Epstein, 1992), one probabilistic choice task (dice problem; Gal & Baron, 1996; West & Stanovich, 2003), one covariation detection task and one gambler’s fallacy task (coin problem; Toplak, West & Stanovich, 2011), the cognitive reflection test (CRT) consisting of three questions (Frederick, 2005), one syllogistic reasoning task (Evans, Barston & Pollard, 1983), and one deductive reasoning task (George, 1995). The applicants had 15 minutes to complete this test and received two movie tickets each for their participation.

We wanted to confirm that there was some diagnostic information available in the resumes, so that the participants would have a fair chance at accurately estimating the applicants’ performance. A regression analysis with SAT score, GPA, perceived skills average, and students’ own guesses of their upcoming performance on the test, confirmed that this was the case. The model turned out significant with R square = 0.73, p = 0.012. Because of the sample being so small, alone none of the individual predictors reached p < 0.05, although the SAT score, which was the strongest predictor, was marginally significant with p = 0.059 (beta = 0.67, 95% CI [-0.19, 8.54]).

Because pilot testing revealed that the task of reading and evaluating all 16 applications was too cognitively demanding, and because we did not want to sacrifice participants’ attention and effort spent on the task, we chose to reduce the number of applicants to eight (four male and four female). Evaluating eight applicants took between 25 to 50 minutes. The performance of these eight students ranged from four to 11 correct answers on the cognitive test. We confirmed that the pattern of variables predicting performance remained similar when reducing the number of applicants.

Participants. One-hundred and sixty-one students participated in Study 1. We excluded one participant for failing to provide estimates of performance for all eight applicants. The final sample (n = 160) consisted of students who were enrolled in either organizational and clinical psychology or personnel and organization science courses (114 women and 45 men, one participant did not report gender; age 19–51, Mage = 25.23, SD = 6.10). The chance of winning movie tickets was offered as compensation for participating in the study.

Materials and procedure. The participants were told that they would be provided with eight applications for a position as research assistant and
that their task was to guess the cognitive performance of each applicant and also rate the extent that they perceived the applicant to be employable for the position. The participants provided informed consent and read a more thorough description of the research assistant position and the cognitive test.

The applicants were presented in the order 1–8 or 5–8, 1–4. Apart from the CV and personal letter, the participants were also presented with the candidate’s subjective estimation of his or her own performance and perceived skills. For each application, the participants provided an estimate of how many questions (0–11) they thought that the applicant had answered correctly on the cognitive test. Next, they were asked how certain they were of this estimation (1 = very uncertain, 5 = very certain). They also rated how qualified (1 = very unqualified, 5 = very qualified), suitable (1 = very unsuitable, 5 = very suitable), and employable (1 = not at all employable, 5 = very employable) they perceived the applicant to be for the job. These three items were averaged to an index with Cronbach’s alpha ranging from 0.85 to 0.90 for the eight applicants.

After evaluating all eight applications, the participants read the following: “Some people prefer to use an immediate feeling that arises spontaneously when forming an impression of a job candidate, while others instead rely more on systematic and analytic thinking when evaluating job candidates. We are interested in which of these two strategies you used when you evaluated the eight candidates. To what extent did you use your intuition/gut feeling compared to analytical thinking when evaluating the applicants?” (1 = only analytical thinking, 5 = only intuition).

Next, the participants rated the extent to which they based their evaluations on the applicants’ personal letter, education, GPA, and SAT scores, respectively (1 = not at all, 5 = completely).

Finally, they provided information about previous experience in working with selection and recruitment (1 = none, 2 = some, 3 = moderate, 4 = substantial, 5 = extensive), their age, and gender.

Data preparations. To calculate accuracy we computed a variable for the mean absolute deviation across all eight applicants so that zero meant an accurate estimation and the larger the deviation from zero, the more inaccurate the estimation of performance, regardless of whether it was under- or overestimation.

As expected, using one’s intuition to a higher extent corresponded to basing one’s impression of the candidate on their personal letter to a higher extent, $r = 0.37, p < 0.001$, and on the SAT equivalent to a lesser extent, $r = −0.16, p = 0.06$ (marginally significant). It also corresponded to incorporating education and grades to a lesser extent; however, these relationships were not significant (both $r = −0.10, p = 0.20$). In sum, we found tentative support for the hypothesis that using one’s intuition would translate into basing one’s impression of the candidates more on the holistic description communicated in the personal letter and less on diagnostic cues. As expected, basing impressions on the SAT equivalent meant more accurate assessments, $r = −0.18, p = 0.02$, confirming that participants who reported using more valid cues also were more accurate.

Contrary to our expectations, individuals who used their intuition were not more confident in their judgments ($r = −0.01, p = 0.91$). Furthermore, confidence did not correlate significantly with accuracy ($r = −0.10, p = 0.20$).

Finally, we confirmed that judgments of applicant cognitive performance predicted perceived applicant employability; the strength of the correlation coefficients was moderate for most of the applicants ($r$s were between 0.25 and 0.62). Having established that students’ thinking mode and attention to specific types of resume content predict accurate judgments of applicant cognitive ability and that these judgments influence perceived employability, we proceeded to examine our main research questions with professional recruiters.

### Study 2

In Study 2, we examined whether there are any differences between experienced recruiters and laypeople in terms of accuracy in predictions of applicant cognitive performance. We also examined whether the degree of feedback on selection judgments that recruiters report to receive predict accuracy. Another aim was to experimentally test whether adopting an analytical versus intuitive mindset matters for accuracy in these judgments. We were also interested in whether experienced recruiters, compared to laypeople, benefit from using their intuition compared to analytic thinking. Because previous research (to our knowledge) on these matters is lacking, we approached these research questions in an exploratory manner.

#### Table 1. Descriptive Statistics and Pearson Correlations for Variables in Study 1

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Accuracy</td>
<td>2.23</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Use of intuition in judgments</td>
<td>3.38</td>
<td>0.38</td>
<td>0.17*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Confidence in judgments</td>
<td>3.07</td>
<td>0.66</td>
<td>−0.10</td>
<td>−0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Based judgments on SAT equivalent</td>
<td>2.82</td>
<td>1.21</td>
<td>−0.18*</td>
<td>−0.16</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Based judgments on grades</td>
<td>2.61</td>
<td>1.02</td>
<td>−0.01</td>
<td>−0.10</td>
<td>−0.02</td>
<td>0.48**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Based judgments on education</td>
<td>3.44</td>
<td>0.86</td>
<td>0.11</td>
<td>−0.10</td>
<td>0.12</td>
<td>0.09</td>
<td>0.24**</td>
<td></td>
</tr>
<tr>
<td>7. Based judgments on personal letter</td>
<td>3.41</td>
<td>0.97</td>
<td>−0.09</td>
<td>0.37**</td>
<td>0.09</td>
<td>−0.20*</td>
<td>−0.28**</td>
<td>−0.24**</td>
</tr>
</tbody>
</table>

Notes: Accuracy was computed as the mean absolute deviation across all eight applicants so that zero meant an accurate estimation and the larger the deviation from zero, the more inaccurate the estimation of performance, regardless of whether it was under- or overestimation. All the variables except accuracy were indicated on a scale from 1 to 5.

*p < 0.05
**p < 0.01, two-tailed.
Past findings suggest that "experts" typically exhibit poor interjudge agreement (Dawes, 1996; Hastie & Dawes, 2001), and that this may also be true for recruiters (Cole et al., 2009). We were therefore interested in whether recruiters’ agreement in their predictions of applicant performance is similar to laypeople’s agreement or if the two groups differ in this regard.

Based on previous research that shows increased confidence following experience with a task (Sanchez & Dunning, 2018), we expected recruiters’ level of experience on the job (as measured with a continuous variable) to correlate positively with confidence in their judgments, and we expected recruiters to display higher confidence overall in their judgments, compared to laypeople. As in Study 1, we also expected individuals who choose to rely more on their intuition to feel more confident in their judgments.

Regarding biographical information, we expected the use of intuition to be associated with more attention to the holistic picture presented in the personal letter, and less attention to the diagnostic cue of SAT scores. We also predicted that relying on SAT scores would correspond to being more accurate. Finally, we expected judgments of applicant cognitive performance to predict judgments of employability.

Method

Participants and design. The participants were 78 laypeople and 87 individuals with recruitment experience, who were randomly assigned to the intuitive thinking (n = 88) or the analytical thinking condition (n = 77).

The recruiters (58 women and 29 men, M(age) = 35.46, SD = 10.04) were contacted primarily via e-mail, and they received a movie ticket for their participation. They rated their experience of working as a recruiter on a seven-point scale (1 = none, 4 = moderate, 7 = extensive, M = 4.75, SD = 1.67). They also answered how often they had received feedback on their recruitment decisions, that is, how often they found out to what extent the recruitment outcome had been successful (1 = never, 4 = sometimes, 7 = very often, M = 4.93, SD = 2.02). Of the recruiters, 39.1% had received no training in recruitment and assessment, 34.5% had undergone university training, and 21.8% had received no training in recruitment and assessment.

The majority of the laypeople were students from Swedish universities (55 women and 23 men, M(age) = 25.40, SD = 6.21). Out of these, 71.8% reported no training in recruitment and assessment, 25.6% had undergone university training, and 1.3% reported other (unspecified) training. Laypeople were offered the chance of winning movie tickets as compensation for their participation.

Materials and procedure. The materials and procedure for evaluations of applicants were identical to Study 1 except for the following. In Study 2 we added an experimental manipulation of information processing strategy, with the design being a 2 (recruiter vs. layperson) × 2 (intuitive vs. analytical thinking) between subjects factorial. After receiving general instructions for the task and giving their informed consent, the participants read the instruction that served as the experimental manipulation. Similar to Rusou, Zakay, and Usher (2013), we used a modified version of Wilson and Schooler’s (1991) manipulation, that contrasts direct choice with a deliberative analysis, and relies on a dual-processing perspective that regards intuitive thinking as the default mode, automatically generating impressions that can then be rejected by conscious effortful reasoning:

When you read the resumes and form an impression of the candidates, it is very important that you rely on your spontaneous intuition/"gut-feeling" about the candidates.

That is, we want you to base your judgment on your immediate impression of the candidate, without overthinking. (Intuitive thinking condition.)

Versus:

When you read the resumes and form an impression of the candidates, it is very important that you think carefully. That is, we want you to analyze the different aspects of the applications thoroughly, without jumping to conclusions. At the end of the survey, you will be asked to respond to questions about how you incorporated these different aspects into your judgments. (Analytical thinking condition.)

To intensify the manipulation, we added a reminder to focus on the immediate gut-feeling (intuitive thinking condition) or to think through and analyze (analytical thinking condition), before the participant provided their impressions for each of the eight applicants.

We used a seven-point scale for all continuous measures, and included a question about thinking mode to serve as a manipulation check: “To what extent did you use your intuition/gut-feeling compared to analytical thinking when judging the applicants?” The recruiters (but not the laypeople) also answered the question “To what extent do you generally use your intuition compared to analytical thinking in recruitment judgments?” (1 = entirely analytical thinking, 7 = entirely intuition).

Data preparations. Descriptive statistics can be found in Table 2 and correlations in Table 3. The (in)accuracy variable was calculated as the mean absolute deviation across all eight applicants (same as in Study 1).

Results

To examine whether thinking mode matters and whether recruiters differ from laypeople with respect to accuracy, we performed an ANOVA with group (recruiters vs. laypeople) and experimental condition (intuition vs. analytical thinking) as independent variables, and deviation from accuracy as the dependent variable. The main effect of group (recruiter vs. laypeople) was not significant and had a very small effect size, F(1, 161) < 0.0005, p = 1.00, η² = 0.0005. The main effect of thinking mode was also very small and statistically non-significant, F(1, 161) = 0.02, p = 0.90, η² < 0.0005. Finally, the interaction effect was small and statistically non-significant, F(1, 161) = 0.12, p = 0.73, η² = 0.001, suggesting that recruiters do not benefit more from using intuitive thinking compared to

| Table 2. Descriptive Statistics for Variables in Study 2 |
|-----------------|-------|-------|-------|-------|
|                 | Recruiters | Laypeople |     |     |
|                 | M      | SD     | M      | SD     |
| Accuracy Use of intuition in judgments (manipulation check) | 2.48 | 0.78 | 2.48 | 0.78 |
| Confidence in judgments | 4.26 | 1.34 | 4.50 | 1.28 |
| Based judgments on SAT equivalent | 4.30 | 1.10 | 4.44 | 0.82 |
| Based judgments on grades | 3.28 | 1.73 | 3.50 | 1.67 |
| Based judgments on education | 2.72 | 1.38 | 3.19 | 1.64 |
| Based judgments on personal letter | 4.33 | 1.29 | 4.36 | 1.25 |

Note: All the variables except for accuracy were indicated on a scale from 1 to 7. Accuracy was computed as the mean absolute deviation across all eight applicants so that zero meant an accurate estimation and the larger the deviation from zero, the more inaccurate the estimation of performance, regardless of whether it was under- or overestimation.
laypeople. Inspection of the manipulation check confirmed that the participants in the intuitive thinking condition reported that they had relied on their intuition more (M = 4.77, SD = 1.15) than participants in the analytical thinking condition (M = 3.92, SD = 1.35), t(163) = −4.38, p < 0.001, 95% CI [−1.24, −0.47], d = 0.68.

Having established that there were very small differences between recruiters and laypeople, we examined whether the two groups differed in the within-group variance of their performance estimations (unrelated to accuracy) for each applicant. A series of Levene’s tests (one for each applicant), with a more stringent alpha level set to p = 0.002 due to multiple tests, showed that none of the eight estimations differed significantly in variance between recruiters and laypeople (ps were between 0.21 and 0.997). Put differently, recruiters did not display higher agreement with each other in their predictions of applicant performance than did laypeople.

Furthermore, we found that the degree of feedback that recruiters reported to receive on their selection judgments in general did not significantly predict accuracy, r = −0.08, p = 0.46. Nor did degree of job experience predict accuracy, r = 0.14, p = 0.19.

As in Study 1, we confirmed that estimations of applicant cognitive performance significantly predicted perceived applicant employability; rs were between 0.45 and 0.72 (recruiters) and between 0.40 and 0.65 (laypeople), for the eight applicants.

In line with our expectations and similar to Study 1, the more the participants claimed that they had used their intuition in the task, the more they reported basing their judgments on the way that applicants introduced themselves in the personal letter, r = 0.26, p = 0.001. The relationship was somewhat stronger for laypeople (r = 0.35, p = 0.02) than for recruiters (r = 0.19, p = 0.08), although this difference between the two groups was not significant according to a Fischer’s z-test of correlation coefficients (p = 0.28, two-tailed; Preacher, 2002). Self-reported use of intuition was not significantly related to any of the other biographical information variables (SAT score, GPA, or education, ps > 0.15), meaning that we did not find support for the hypothesis that intuition would correspond to reduced reliance on diagnostic cues.

Recruiters and laypeople did not differ significantly in their reliance on applicants’ education (p = 0.90) or SAT scores (p = 0.40), but laypeople (M = 3.19, SD = 1.64) relied on GPA to a higher extent than recruiters (M = 2.72, SD = 1.38), t(163) = 1.99, p = 0.048, 95% CI [0.003, 0.93], d = 0.31. They also relied on the personal letter (M = 4.76, SD = 1.25) more than recruiters did (M = 4.25, SD = 1.60), t(163) = 2.23, p = 0.03, 95% CI [0.06, 0.95], d = 0.36.

As in Study 1, we found that basing one’s impressions of applicant cognitive performance on SAT scores meant more accurate assessments, r = −0.18, p = 0.02, confirming that participants who reported using the most valid cue were more accurate. There were no other significant associations between basing judgments on specific resume content and accuracy (ps > 0.21).

Contrary to our expectations, the two groups did not differ significantly in how confident they felt about their judgments (Mrecruiters = 4.30, SD = 1.09, Mlaypeople = 4.44, SD = 0.82), t (163) = 0.92, p = 0.36, 95% CI [-0.16 - 0.44], d = 0.15. Also in contrast to our expectations, recruiters’ amount of experience on the job was not significantly related to confidence, r = 0.10, p = 0.38. Further, there was a weak association between higher confidence and more accurate judgments (r = −0.13, p = 0.09).

Finally, we found that the expected association between using one’s intuition to a higher extent when making the judgments and higher confidence was not significant for recruiters, r = 0.08, p = 0.45, but it was for laypeople, r = 0.25, p = 0.03.

In sum, there were no apparent differences between recruiters and laypeople in terms of accuracy or agreement, and neither recruiters’ degree of feedback nor degree of job experience significantly predicted accuracy. Despite a significant manipulation check, there were no effects of experimentally induced thinking mode on accuracy.

The hypothesis that estimations of applicant performance would predict perceived applicant employability received clear support. The hypothesis that self-reported use of intuition would predict use of certain biographical information received only partial support, but it was confirmed that those who reported using the most valid cue (SAT scores) were more accurate.

Table 3. Pearson Correlations for Variables in Study 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Accuracy</td>
<td>–</td>
<td>0.14</td>
<td>–0.08</td>
<td>0.14</td>
<td>–0.08</td>
<td>–0.30**</td>
<td>–0.21</td>
<td>0.17</td>
<td>–0.13</td>
</tr>
<tr>
<td>2. Job experience</td>
<td>n/a</td>
<td>n/a</td>
<td>0.64**</td>
<td>–0.03</td>
<td>0.10</td>
<td>–0.08</td>
<td>–0.15</td>
<td>–0.22*</td>
<td>–0.08</td>
</tr>
<tr>
<td>3. Feedback</td>
<td>n/a</td>
<td>n/a</td>
<td>–</td>
<td>–0.05</td>
<td>0.26*</td>
<td>0.07</td>
<td>–0.02</td>
<td>–0.24*</td>
<td>–0.11</td>
</tr>
<tr>
<td>4. Use of intuition in judgments (manipulation check)</td>
<td>–0.10</td>
<td>n/a</td>
<td>n/a</td>
<td>0.08</td>
<td>–0.02</td>
<td>–0.02</td>
<td>–0.09</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>5. Confidence in judgments</td>
<td>–0.22</td>
<td>n/a</td>
<td>n/a</td>
<td>0.25*</td>
<td>–</td>
<td>0.01</td>
<td>0.16</td>
<td>0.06</td>
<td>0.18</td>
</tr>
<tr>
<td>6. Based judgments on SAT equivalent</td>
<td>–0.03</td>
<td>n/a</td>
<td>n/a</td>
<td>–0.02</td>
<td>0.08</td>
<td>–</td>
<td>0.60**</td>
<td>–0.06</td>
<td>–0.12</td>
</tr>
<tr>
<td>7. Based judgments on grades</td>
<td>0.06</td>
<td>n/a</td>
<td>n/a</td>
<td>–0.01</td>
<td>0.17</td>
<td>0.67**</td>
<td>–</td>
<td>0.03</td>
<td>0.13</td>
</tr>
<tr>
<td>8. Based judgments on education</td>
<td>–0.15</td>
<td>n/a</td>
<td>n/a</td>
<td>–0.15</td>
<td>0.08</td>
<td>0.13</td>
<td>0.15</td>
<td>–</td>
<td>0.18</td>
</tr>
<tr>
<td>9. Based judgments on personal letter</td>
<td>–0.06</td>
<td>n/a</td>
<td>n/a</td>
<td>0.35**</td>
<td>0.26*</td>
<td>–0.04</td>
<td>0.10</td>
<td>–0.21</td>
<td>–</td>
</tr>
</tbody>
</table>

Notes: The results for recruiters (n = 87) are shown above the diagonal. The results for laypeople (n = 78) are shown below the diagonal. Accuracy was computed as the mean absolute deviation across all eight applicants so that zero meant an accurate estimation and the larger the deviation from zero, the more inaccurate the estimation of performance, regardless of whether it was under- or overestimation.

*p < 0.05
**p < 0.01, two-tailed.
Finally, the hypothesized relationships with confidence received weak support overall.

GENERAL DISCUSSION

The point of spending resources and time on the selection process is to predict which job candidate who is most likely to perform well on the job. Cognitive ability is among the top predictors of future work performance (Robertson & Smith, 2001; Schmidt & Hunter, 1998) and is frequently used in candidate evaluation (Huffcutt et al., 2001). Although it can be assessed through cognitive ability tests, it is common practice for recruiters to base their impression on applicants’ resumes. In many cases, resume-based impressions of the applicant’s competence (including cognitive ability) therefore have a direct impact on the final selection decision. The current research is, to our knowledge, the first to investigate accuracy in recruiters’ compared to laypeople’s predictions of cognitive performance. It makes an important contribution to the literature by providing preliminary evidence on the role of expertise and thinking mode for accuracy in these judgments.

The results revealed that recruiters did not outperform laypeople, and that neither degree of work experience or degree of previously received feedback were significantly associated with accuracy. Akin to previous research on judgments of personality (Schmid Mast et al., 2011) and leadership ability (Re & Rule, 2016), our findings thus clearly suggest that experience on the job is not a guarantee of success when making judgments of applicant cognitive ability. Similar to Cole et al. (2009) who found low inter-rater reliability in recruiters’ judgments of applicant personality from resumes, we found that recruiters did not display more agreement with each other in their judgments than laypeople. This suggests that some recruiters, despite having moderate experience, may not have developed more systematic ways (compared to laypeople) of integrating multiple sources of information prior to making judgments of job candidates’ abilities. To improve predictive power in real selection situations, it might help if judgments are averaged across multiple raters (Kuncel et al., 2013).

Resumes provide a tool to become acquainted with the applicants and is a cheap and easy means of accomplishing an initial round of screening. Nonetheless, our findings call for a word of caution as recruiters cannot simply be assumed to possess superior expertise (relative to laypeople) in the art of screening resumes for competent applicants. A safer approach when trying to get an idea of applicants’ cognitive ability would probably be to use cognitive ability tests, even though a belief in “the myth of selection expertise” can make many recruiters reluctant to use valid selection aids (Highhouse, 2008).

Our finding that recruiters did not benefit from using their intuition suggests that they did not possess intuitive expertise. To answer the question of why recruiters do not outperform laypeople and why they may not benefit from using their intuition, it might be helpful to look at their typical learning environment. Because of a delay between selection decisions and finding out whether the selected applicant eventually performs well on the job (that is, if the recruiter finds out at all), recruiter’s work conditions are probably not optimal for developing a reliable expert intuition. Put differently, it seems likely that their work situation does not provide opportunities for regular, immediate, and unambiguous feedback on judgments of applicants’ attributes, which is often necessary for intuitive expertise to develop (Kahneman & Klein, 2009).

On a positive note, if intuition is largely the result of learning (Hogarth, 2010), this implies that intuition can be explicitly educated because doing so would be similar to defining the conditions that allow task-specific expertise to develop. Indeed, most people can be trained to develop high levels of accuracy for many tasks if they are exposed to many forced choice trials with feedback (Myers, 2002; Seligman & Kahana, 2009), suggesting the possibility that this may also be the case for selection judgments. For such education to be successful, experience would have to be organized so that recruiters learn the right lessons from their interactions with applicants and from resumes. Importantly, improvements from training may require that individuals receive clear feedback before they internalize the use of certain cues (Re & Rule, 2016), suggesting that people should undergo such training before they start working with selection and recruitment. Taken together, it may very well be the case that intuitive expertise in a personnel selection context is possible to develop, but that this requires time and effort. We suggest that more research is needed on the role of feedback, preferably using a longitudinal design where type and degree of feedback is experimentally manipulated. For now, the results of Study 1 tentatively suggest that intuition should be avoided when it comes to resume screeners with low experience. Nonetheless, our findings do not allow for conclusions beyond judgments of cognitive ability, and it is of course possible that intuition will work better when predicting other qualities that are important for work adjustment and success in certain occupations, such as for example social skills.

Although there were some differences in the weight that recruiters and laypeople put on different resume content, they did not differ in their use of the most valid cue (SAT scores). Because some types of biographical information predict applicant cognitive ability, and participants who relied on more diagnostic cues were more accurate, it may be feasible to train recruiters in detecting and assessing specific cues instead of using unsystematic reviews of resume content to make far-reaching generalizations about applicants (Cole et al., 2003). Consistent with this suggestion, recent work with student raters in a laboratory environment showed that increased systematicity leads to selection of more qualified candidates (Bäckström & Björklund, 2017). Moreover, Scholz, von Helversen, and Rieskamp (2015) found that individuals using rule-based decision-making when evaluating job applicants avoid relying on previously seen exemplars, such as recently encountered applicants.

Similar to previous work that has found subjective impressions based on resumes to influence employability judgments (Burns et al., 2014; Cable & Gilovich, 1998; Cole et al., 2009), we found that impressions of cognitive ability influenced judgments of applicant employability. This implies that if judgments of applicant competence are inaccurate, these judgments can have detrimental consequences for organizations, such as failing to select the most competent candidates.
In the current research, we also examined how confidence in judgments of applicant cognitive ability relates to expertise and thinking mode. Its relation to intuition revealed mixed results. For laypeople, but not for recruiters or students, feeling more confident was associated with greater reliance on intuition. Moreover, we found weak and non-significant associations between confidence and accuracy, which is similar to the Schmid Mast et al. (2011) study where recruiters and laypeople’s self-evaluated assessment skills were not significantly associated with their personality assessment performance (although they found a weak association with deception detection). Furthermore, although we expected experience in working with personnel selection to predict confidence judgments, we found no support for this hypothesis. Investigations on the role of confidence in selection research are scarce (Kausel, Culbertson & Madrid, 2016) and more research is needed to develop this line of inquiry further.

Strengths, limitations, and further directions for research

The research reported in this paper has its strengths as well as weaknesses. First, in contrast to most studies on bias and accuracy in resume screening (e.g., Cole et al., 2005; Watkins & Johnston, 2000), we tested our main research questions with real recruiters. Second, our applicants were real people who provided their own personal biographical information. Third, we tried to reduce the risk of common method bias by measuring judgments of applicant performance on the cognitive ability test and employability ratings with different scales (as recommended by Podsakoff, MacKenzie, Lee & Podsakoff, 2003; also see Burns et al., 2014). This suggests that there was some similarity to real resume screening settings. Nonetheless, we cannot rule out that this situation differs from a real recruitment situation, and follow-up research which examines the judgment processes in high-stake selection situations would be optimal. Furthermore, even though our pilot study revealed that the resume content contained information that was diagnostic of the applicants’ performance on the cognitive test, it is possible that the resumes that we used are not perfectly representative of the broader population of job applicant resumes, and preferably, our findings should thus be replicated with a larger set of applicant stimuli.

Another limitation pertaining to generalizability is that our sample of recruiters consisted primarily of individuals with a moderate level of experience in working with personnel selection, suggesting that they should perhaps be classified as experienced professionals rather than experts. Because previous research suggests that individuals with mediocre level of experience sometimes process information differently from accomplished experts (Baylor, 2001), some caution should be exercised when drawing conclusions about intuitive expertise from our findings (Sinclair, 2010). More specifically, we cannot be certain of the extent to which our findings will generalize to recruiters with very high expertise. After all, it is possible that recruiters with very high expertise would indeed benefit from using their intuition. It is also possible that such recruiters are more accurate in their judgments than recruiters with moderate work experience. Ideally, future research should address this question. However, we still feel confident that our results can be generalized to the majority of recruiters, as most recruiters are unlikely to be extremely high in expertise for the reasons outlined above.

In Study 2, we experimentally manipulated whether the participants were instructed to use an intuitive or analytical approach. Although our design allowed for an adequate test of our hypothesis concerning the relative usefulness of relying on intuition versus analytical thinking, it did not include a baseline (no instruction) condition to which one could compare the effects of different thinking modes. This could be seen as a limitation. Moreover, although a significant manipulation check suggested that the manipulation worked as intended, future studies may want to develop and subsequently validate other experimental manipulations of intuition and analytical thinking.

On a final note, there are several ways to examine accuracy and bias in judgment and decision-making. In the present research, we examined whether there were any differences in accuracy relative to other variables. As we did not focus on accuracy in absolute terms, we did not examine whether there were any differences between recruiters and laypeople in inappropriately elevated or depressed ratings. Future research might want to pursue to what degree inaccurate judgments reflect overestimation versus underestimation.

The present research attests to the difficulty inherent in the task faced by recruiters when screening resumes to find the most competent applicants. We hope that this work will spawn further research efforts on this important topic so that selection judgments can become more accurate and decisions more effective. After all, this would not only be in the best interest of organizations, but job applicants as well.

AUTHOR CONTRIBUTIONS

JA developed the overall study concept. SS planned and designed the individual studies with help from JA. SS collected the data for the pilot study and Study 1, and supervised the data collection for Study 2. SS conducted the statistical analyses and wrote the manuscript. JA provided critical comments on the manuscript. Both authors approved the final version for submission.

The authors thank Tara Said for assistance with data collection as part of her master’s thesis at Linnaeus University.

FUNDING

This research was supported by grant 20160536 from the Crafoord Foundation awarded to Jens Agerström.

REFERENCES


© 2020 The Authors. *Scandinavian Journal of Psychology* published by Scandinavian Psychological Associations and John Wiley & Sons Ltd