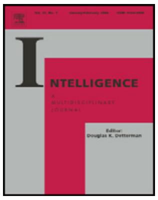
ARTICLE IN PRESS
INTELL-00641; No of Pages 12
Intelligence xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Intelligence



General intelligence predicts reasoning ability even for evolutionarily familiar content

Scott Barry Kaufman [a,*,1], Colin G. DeYoung [b], Deidre L. Reis [c,d], Jeremy R. Gray [c,e]

[a] New York University, Department of Psychology, USA
[b] University of Minnesota, Department of Psychology, USA
[c] Yale University, Department of Psychology, USA
[d] Hunter College, Department of Psychology, USA
[e] Yale University, Interdepartmental Neuroscience Program, USA

ARTICLE INFO

Article history:
Received 7 January 2011
Received in revised form 12 May 2011
Accepted 17 May 2011
Available online xxxx

Keywords:
Intelligence
g factor
Evolutionary psychology
Deductive reasoning
Wason Card Selection Task

ABSTRACT

The existence of general-purpose cognitive mechanisms related to intelligence, which appear to facilitate all forms of problem solving, conflicts with the strong modularity view of the mind espoused by some evolutionary psychologists. The current study assessed the contribution of general intelligence (*g*) to explaining variation in contextualized deductive reasoning. One hundred and twelve participants solved 70 contextualized reasoning problems in a computerized version of the Wason Card Selection Task that recorded both accuracy and reaction time. Consistent with prior research, in the sample as a whole, precautionary and social exchange reasoning problems were solved more frequently and more quickly than reasoning problems about arbitrary rules. At the individual-differences level of analysis, however, performance on all reasoning tests was significantly correlated and loaded on a single deductive-reasoning accuracy factor. Further, this factor was significantly correlated with *g*. There was no relation, however, between *g* and the speed of arriving at the correct answer for any form of deductive reasoning. We discuss the implications of these findings for evolutionary psychology, intelligence, and reasoning.

1. The role of general intelligence in contextualized deductive reasoning

Over 100 years ago, Spearman (1904) first discovered a consistent tendency for a diverse range of cognitive tests with differing content to be positively correlated with one another, a phenomenon he described as the “positive manifold,” which suggests the existence of a general intelligence factor (*g*). The existence of the *g* factor is such a robust finding that one should expect performance on any reasonably complex explicit cognitive task (in contrast to implicit cognitive tasks such as implicit learning) to be associated with *g* (Carroll, 1993; Chabris, 2007; Gottfredson, 2002; Jensen, 1998; Johnson & Bouchard, 2005). A standard hypothesis in factor analysis is that the variables loading on a single factor co-vary due to a shared underlying cause or set of causes (Haig, 2005; although this is not a necessary condition for the existence of a factor; Bartholomew, Deary, & Lawn, 2009). The existence of *g* thus suggests the possibility of causal forces that influence performance on most complex cognitive tasks. These forces might be of at least two kinds: (1) problems such as genetic mutations or developmental abnormalities that influence many different cognitive mechanisms (Keller & Miller, 2006; Arden, Gottfredson, Miller, & Pierce, 2009; Yeo, Gangestad, Liu, Calhoun, & Hutchison, 2011); (2) cognitive mechanisms that are utilized to some extent in most or all complex

---

* Corresponding author at: Department of Psychology, New York University, 6 Washington Place, room 158, New York, New York, 10003, USA.
E-mail address: scott.barry.kaufman@nyu.edu (S.B. Kaufman).

[1] The first author conducted this study in Cambridge, England while he was a doctoral student at Yale University.

0160-2896/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2011.05.002

Please cite this article as: Kaufman, S.B., et al., General intelligence predicts reasoning ability even for evolutionarily familiar content, *Intelligence* (2011), doi:10.1016/j.intell.2011.05.002

cognitive tasks. In the present study we focus on the second of these possibilities, but note that the two are not mutually exclusive. Psychological and neural evidence suggests that *g* is not caused by a single unitary process, but is instead the result of multiple cognitive mechanisms (Jung & Haier, 2007; Kaufman, DeYoung, Gray, Brown, & Mackintosh, 2009; van der Maas et al., 2006).

Constructs, like *g*, that describe dimensions of variation in a population are neither identical to nor necessitate constructs, like cognitive mechanisms, that refer to processes in an individual. However, individual-differences research can provide relevant evidence to link the two types of construct using the principal that if process X is involved in trait Y, then individual differences in X should predict individual differences in Y (though, of course, the correlation of X and Y does not provide evidence that X causes Y). Some evidence exists to suggest that cognitive processes like working memory, explicit associative learning, and information processing speed are cognitive mechanisms involved in general intelligence (Kaufman et al., 2009). Nonetheless, the existence of cognitive mechanisms for general intelligence is controversial in evolutionary psychology. A central tenet of evolutionary psychology is that natural selection sculpted the human mind to solve specific recurring problems of survival and reproduction, and that therefore the mind consists of multiple domain-specific mental mechanisms or “modules” that are activated by specific contexts (Cosmides & Tooby, 2001). At first blush, the existence of a general factor of intelligence might appear incompatible with a strong modularity view of human cognition because *g* is domain-general rather than domain-specific: it is associated with performance on cognitive tasks in a multitude of different contexts. Although a number of evolutionary psychologists have acknowledged the existence of domain-general cognitive processes, and some have explicitly related them to *g* (e.g., Chiappe & MacDonald, 2005; Deaner, van Schaik, & Johnson, 2006; Geary, 2004, 2009; Penke, 2010; Sperber, 1994), evolutionary psychologists often downplay their importance relative to domain-specific modules (Cosmides, Barrett, & Tooby, 2010; Cosmides & Tooby, 2001).

In the present study, we explored this tension between evolutionary psychology and the theory of general intelligence by examining individual differences in a cognitive paradigm that has been used extensively by evolutionary psychologists to provide evidence that cognitive abilities are domain-specific rather than domain-general: the Wason four-card selection task (Wason, 1968). In the Wason task, participants consider a rule of the generic form “If P then Q” along with four cards describing P or not-P on one side and Q or not-Q on the other (with one of each type face up). The participant is told to indicate only the cards that definitely must be turned over to determine if the rule is being broken (correct answer: P and not-Q).

The selection task is a useful tool to investigate the nature of human cognition because it is sensitive to the content and context of presentation (Evans, 2003): performance is typically poor when rules are decontextualized, abstract, or arbitrary in nature, but often quite good when problems involve potential transgressions of social norms or precautionary reasoning about physically dangerous situations (Cosmides, 1989; Fiddick, Cosmides, & Tooby, 2000; Gigerenzer & Hug, 1992).

Contextualized deductive reasoning involves if-then reasoning put into a narrative vignette context. “Social exchange” problems concern the mutual exchange of goods or services between individuals in specific situations. The rules generally involve detecting if one party might be taking a benefit without fulfilling an obligation (e.g., “If you borrow my motorcycle, then you have to wash it.”). “Precautionary” problems involve rules related to avoiding potential physical danger (e.g., “If you surf in cold water, then you have to wear a wetsuit”). “Arbitrary-rule” problems (Cheng & Holyoak, 1985) have arbitrary rules (e.g., “If the soda is diet, then it has to be in a purple container”) that are nonetheless contextualized in realistic scenarios. Note that by “arbitrary,” we are not referring to cultural rules that are evolutionarily arbitrary but might nonetheless be influenced by evolved heuristics regarding social exchange or social norms (e.g., men must take off their hats indoors, but women may leave them on). Rather, these rules are arbitrary in that they do not correspond to established rules in the individual's experience.

The difference in performance on precautionary and social exchange vs. arbitrary-rule problems has been explained with reference to the concept of modularity (Cosmides & Tooby, 2004; Cosmides, 1989). Reasoning about social exchange and precautions is hypothesized to be supported by dedicated information processing modules that result from evolutionary selection pressure exerted by situations involving social exchange or physical danger. No such pressure has been exerted by situations involving the arbitrary rules, and thus the human mind does not have a cognitive module that allows accurate reasoning about arbitrary rules.

Note that this hypothesis describes the cognitive functions of humans as a group, and says nothing regarding individual differences. Indeed, evolutionary psychology has generally had little to say about individual differences, in part because of the assumption that, for any trait important to fitness, selection pressure would reduce variance around an optimal level of the trait, with individual differences being mere random noise. Recently, however, more attention has been paid to evolutionary processes that would maintain variation in traits that do have adaptive significance, including fluctuating selection (in which higher levels of a trait are more adaptive in some environments and lower levels are more adaptive in others) and the difficulty of maintaining certain traits in the face of factors (such as deleterious mutations) that reduce fitness (Arden et al., 2008; Buss & Hawley, 2010; Furlow, Armijo-Prewitt, Gangestad, & Thornhill, 1997; Prokosch, Yeo, & Miller, 2005; ; Keller & Miller, 2006; Miller, 2000; Nettle, 2006).

Intelligence has been proposed as an example of the latter process, which would make it a fitness indicator: higher intelligence would almost always be associated with increased fitness, but the biological difficulty of producing an individual with high intelligence would ensure that the population maintains a range of intelligence over time, despite selection pressure (Miller, 2000). These ideas open the door to reconciling the existence of general intelligence, both as a set of domain general cognitive mechanisms and as a trait with meaningful individual differences, with evolutionary psychology.

A number of researchers have attempted to unite evolutionary psychology with differential psychology (e.g., Penke, 2010; Kanazawa, 2010). Kanazawa (2004; 2010, but

Please cite this article as: Kaufman, S.B., et al., General intelligence predicts reasoning ability even for evolutionarily familiar content, *Intelligence* (2011), doi:10.1016/j.intell.2011.05.002

see Penke et al., in press) has suggested that general intelligence might be a domain-specific adaptation to evolutionarily novel situations: rare occurrences (such as a fire started by lightning or an unusually severe drought) for which there was no existing adaptation. On this basis, Kanazawa has hypothesized that individual differences in *g* should be related to the ability to solve evolutionarily novel problems but not evolutionarily familiar problems. In relation to the Wason selection task, Kanazawa's hypothesis indicates that *g* should be associated with performance on problems with abstract or arbitrary rules, but not on problems with rules reflecting evolutionarily familiar situations (such as social exchange or precaution) for which dedicated modules could have evolved. This hypothesis has not been well supported by the small literature on individual differences in performance on the Wason task. In a study conducted by Girotto and Tentori (2008), all participants who solved the abstract version of the selection task also solved the social contract version, whereas only 28% of their sample who failed to solve the abstract version also solved the social contract version. Additionally, studies reporting correlations with general cognitive ability have been inconsistent, with one finding weaker correlations of intelligence with performance on nonarbitrary relative to arbitrary problems (Stanovich & West, 1998), another finding just the opposite (Klaczynski, 2001), and a third finding intelligence to be similarly associated with both types of problem (Dominowski & Dallob, 1995).

More recently, Reis et al. (2007) developed a computerized, time-limited, card-by-card presentation version of the Wason Card Selection Task. They found that reasoning about social exchanges and reasoning about precautionary situations were very highly correlated with each other, $r = .87$, even though emotional intelligence predicted social exchange reasoning after controlling for precautionary reasoning, and harm avoidance predicted precautionary reasoning after controlling for social exchange reasoning. Their results suggest that there may indeed be substantial common variance among multiple contextualized deductive reasoning types even if there is also some residual domain-specificity for each type.

The present study improves on the prior literature on the domain-general cognitive correlates of contextualized deductive reasoning in several ways: first, by utilizing the psychometrically improved version of the Wason selection task developed by Reis et al. (2007), which is better suited to assessing individual differences; second, by examining a deductive reasoning factor, based on shared variance across performance on different Wason problems; third, by including a more thorough assessment of *g* that allows creation of factor scores for *g*, and fourth, by including measures of cognitive mechanisms associated with *g*, including working memory, explicit associative learning, and processing speed. These psychometric improvements allow us to perform a more stringent test of Kanazawa's hypothesis that performance on evolutionarily unfamiliar, but not evolutionarily familiar, problems should be associated with general intelligence.

Additionally, we propose our own alternative hypothesis. Kanazawa's hypothesis strikes us as unlikely, both because of the pervasiveness of the positive manifold in cognitive testing and because the logic of his evolutionary argument is debatable. Evolutionarily novel events of the kind that Kanazawa describes are rare by definition. Although rare events can have consequences for evolution if they affect sufficiently large numbers of a species (Kanazawa, 2010), most rare events are likely to affect a small proportion of individuals, and their rarity will prevent them from exerting consistent selection pressure. Therefore, it seems more likely to us that mechanisms for general intelligence would have evolved in response to all situations for which a pre-existing adaptation did not produce an optimal response (cf. Chiappe & MacDonald, 2005; Geary, 2004, 2009; Penke, 2010; Woodley, 2010). Certainly this class of situations would include evolutionarily novel situations of the Kanazawa type; however, it would also include evolutionarily familiar situations of sufficient complexity to interfere with the heuristic response of a dedicated cognitive module or to render its effectiveness uncertain.

Thus, rare evolutionarily novel events are simply one example of a larger class of situations, namely those that are complex and unpredictable (Peterson, 1999; Peterson & Flanders, 2002). Gottfredson (2002) has provided evidence that *g* represents individual differences in the ability to deal with complexity, in everyday life as well as in IQ tests. Increased social group size and rapidly increasing cultural complexity are likely to have rendered pre-existing heuristic adaptations increasingly fallible in human ancestors, thus increasing the selection pressure on domain-general mechanisms that could logically analyze the causal structure of situations even when it was too complex to be adequately processed by modular heuristics. (In this context it is worth noting the correlation of expansion of the neocortex with social group size across primate species; Dunbar & Shultz, 2007; and the corresponding correlation across these species of brain size with domain general learning ability; Deaner, Isler, Burkart, & van Schaik, 2007.) Note that, given this explanation for the evolution of general intelligence, one might argue that it is indeed specific to a given domain (complex and unpredictable situations), but that this domain is very broad (cf. Barrett, 2009).

Evolutionary psychologists sometimes argue that a class of situations must be relatively narrow to exert consistent selection pressure, but this claim is insufficiently justified. Any regularity in the environment can exert selection pressure if it poses a challenge or opportunity to the organism, and whether this will prompt adaptation simply reflects the likelihood that genetic variation might lead to variation in the ability to meet the challenge or seize the opportunity. In the case of complex, unpredictable situations, regardless of their superficial dissimilarity, selection for increased ability to analyze causal structure is highly likely. The mechanisms underlying general intelligence should, therefore, be brought to bear in attempting to solve any explicit cognitively complex problem (Girotto & Tentori, 2008). Existing adaptations may facilitate performance on evolutionarily familiar problems, but general intelligence should provide additional facilitation.

Note that we are not arguing against the existence of cognitive modules, but rather asserting that they are likely to coexist, and to function simultaneously, with important domain-general mechanisms (for similar ideas see Cosmides et al., 2010, and Penke, 2010). Several neuroimaging studies using the

Wason selection task have demonstrated that different brain regions are involved in different types of reasoning, a finding which supports the existence of mechanisms specialized for different types of reasoning (Ermer, Guerin, Cosmides, Tooby, & Miller, 2006; Fiddick, Spampinato, & Grafman, 2005; Reis et al., 2007). However, these results do not contraindicate the additional existence of other mechanisms that are involved in all types of reasoning. An analogous case is that of working memory; different brain systems are involved in spatial vs. verbal working memory, but there are also brain systems that are involved in all types of working memory (Wager & Smith, 2003). Information processing is typically accomplished through a combination of domain-general and domain-specific mechanisms. We therefore hypothesized that, although there would be group-level differences in performance between arbitrary and nonarbitrary problems, nonetheless performance on all types of problems would be correlated with each other and also correlated with *g*, reflecting the additional effect of domain general processes, over and above any species-typical biases conferred by evolved modules.

2. Methods

2.1. Participants

One hundred twelve participants (40 males, 72 females) were included in the analyses presented here and were part of a study that involved 177 participants (Kaufman et al., 2009, 2010). All participants were aged 16–18 years, and attended a selective Sixth Form College (which takes high-achieving students who are in their last 2 years of secondary education) in Cambridge, England. The Wason selection task was offered as an option at the end of the third test session for participants who completed all the other tests and had time remaining. No variables reported here differed significantly between those who did and did not complete the computerized Wason Card Selection Task.

2.2. Computerized Wason Card Selection Task

The computerized version of the Wason Card Selection Task was a shortened version of the task administered by Reis et al. (2007). Both the Reis et al. (2007) version and the shortened version used in the current study are time-limited, involve a card-by-card presentation, and record both accuracy and reaction time. The benefits of administering this version of the Wason Card Selection Task for the current purposes are threefold. Firstly, the card-by-card presentation allows for an assessment of reaction time for each card. Prior research on the Wason selection task has mostly assessed accuracy, whereas the results of Reis et al. (2007) suggest that emotional processes may play an important role in speeding up responses by facilitating reasoning using contextual information. Therefore, the task allows for an assessment of the differing role of cognitive mechanisms in predicting speed vs. accuracy of contextualized deductive reasoning. Secondly, the task assesses multiple forms of contextualized reasoning that have been employed in the literature, allowing for a proper assessment of the common variance across multiple types of contextualized deductive reasoning. The administration of only contextualized reasoning problems (omitting purely abstract, decontextualized problems based on symbols rather than scenarios) allowed us to vary evolutionary relevance without the confound of varying imaginability. Thirdly, prior research on individual differences in performance on the selection task has typically administered relatively few items, which would tend to lead to inadequate psychometric reliability. The current study overcame this limitation by adapting the Wason Card Selection Task developed by Reis et al. (2007), which allows for an assessment of a much larger number of items and therefore an assessment of reliability.

The task involves deductive reasoning on three types of content: arbitrary, precautionary, and social exchange.

Here is an example of a social exchange scenario:

> Joe often goes out to dinner with friends from work, and they go to a bar afterwards. Joe always pays the dinner check with his credit card and his friends pay him back with cash. He notices that people usually don't consider how much their beer costs when paying him back. So Joe announces a rule, “If you order beer at dinner, then you have to buy me a drink at the bar.”
>
> You want to see whether any of Joe's friends cheat on this rule.
>
> The following cards represent five of Joe's friends that joined him for dinner. Each card represents one friend. One side of the card tells what type of drink that person ordered at dinner and the other side tells whether that person bought Joe a drink at the bar.
>
> Please decide if you would definitely need to turn each card over to see if any of the friends cheated on the rule:
>
> “If you order beer at dinner, then you have to buy me a drink at the bar.”
>
> Do not turn over any more cards than are absolutely necessary.

For both arbitrary and social exchange reasoning, participants completed 25 problems in total, 5 items per scenario. Due to a computer malfunction in the recording of responses for one scenario, only 20 problems (5 per scenario) were analyzed for the precautionary trials. For each problem, participants read a brief scenario describing both a situation and a rule of generic form “If P, then must Q.”. They then saw cards presented individually along with the rule. For each scenario, there were one rule and four different cards: P, not-P, Q, not-Q.) For each card, participants chose either “definitely turn over” (correct for P and not-Q) or “no need to turn over” (correct for not-P and Q) to be able to tell whether the rule was being broken. Five cards were used for each scenario, with type of the fifth card (P, not-P, Q, not-Q) varying, to ensure participants could not use process of elimination to figure out the answer to the fourth card since there were only four types of cards. Participants had 20 s to read each scenario and 4 s to respond to each individual card. If a response was not made within 4 s, an error was scored

Please cite this article as: Kaufman, S.B., et al., General intelligence predicts reasoning ability even for evolutionarily familiar content, *Intelligence* (2011), doi:10.1016/j.intell.2011.05.002

and the computer automatically advanced to the next card. No feedback was given about performance.

Scenarios were pseudorandomly ordered and no scenarios were repeated. All participants then received the same scenarios in the same order. To avoid confounding RT with time spent reading each card, the length of text shown on the cards was matched across all three conditions.

Accuracy was calculated by taking the percentage of individual cards responded to correctly for each form of reasoning. RT for each card was measured as the time elapsed from when the card was displayed until the response was made. RT for each problem type (arbitrary, precautionary, social exchange) was computed as the mean of the RTs for all cards responded to correctly, regardless of card type (P, not P, q, not Q). All responses <200 ms were trimmed from all analyses and all timed out trials (>4000 ms) were excluded from the RT analyses.

2.3. General intelligence

To create a well-balanced *g* factor score, we used 6 markers, tapping different domains of cognition. Using one of the largest batteries of cognitive tests ever collected, Johnson and Bouchard (2005) demonstrated that, below the *g* factor, there are three separable second-stratum domains of cognitive ability: verbal, perceptual, and mental rotation. Use of one test from each domain should produce a well-balanced *g*. Additionally, we included three cognitive variables that have been found to contribute unique variance to *g*: explicit associative learning, working memory, and processing speed (Kaufman et al., 2009).

2.3.1. Raven's Advanced Progressive Matrices Test, Set II (APM)

The APM (Raven, Raven, & Court, 1998) is a measure of abstract perceptual reasoning. Each item consists of a 3×3 matrix of geometric patterns with the bottom right pattern missing. The participants' task is to select the option that correctly completes the matrix. There are eight alternative answers for each item. The test is presented in increasing order of difficulty. After two practice items with feedback, participants were then given 45 min to complete 36 items. Descriptive statistics of the APM (M = 22.04, S.D. = 5.76, Range = 7–33) suggested our sample is comparable in IQ to the average undergraduate student (Raven et al., 1998).

2.3.2. DAT verbal reasoning test

The verbal reasoning section of the Differential Aptitudes Test (DAT-V, The Psychological Corporation, 1995) was administered to each participant. Each problem consisted of a sentence with two words missing, and participants chose a pair of words from the answer options that were related to the words in the sentence in some way. After two practice items, participants had 15 min to complete 40 problems.

2.3.3. Mental Rotations Test, Set A (MRT-A)

The MRT-A (Vandenberg & Kruse, 1978) contains 24 problems and measures mental rotation ability, which appears to be a distinct component of intelligence at the same level as verbal ability and perceptual ability (Johnson & Bouchard, 2005). Each problem in the MRT-A shows a three-dimensional target figure paired with four choice figures, two of which are rotated versions of the target figure. To score a point, both rotated versions must be identified. After two practice items with feedback and an explanation, the first 12 problems were attempted in 4 min with a 2 minute break before attempting the second 12 in another 4 min. The maximum score is 24.

2.3.4. Explicit associative learning tasks

2.3.4.1. Three-Term Contingency Learning (Williams & Pearlberg, 2006). The Three-Term Contingency Learning (3-Term) task consists of four learning blocks, each followed immediately by a test block. In each learning block, participants were presented with 10 unique words. Each word was associated with three different words, contingent on a key press. The participants' task was to learn the word associated with each stimulus-response pair. For instance, on one trial the word "LAB" might show on the screen with the letters "A", "B", and "C" listed underneath. When participants selected "A", they saw one association (e.g., PUN), when they selected "B", they saw a second association (e.g., TRY), and when they selected "C" they saw a third association (e.g., EGG). The duration of exposure to each association was self-paced (max 2.5 s) with changeover intervals set at 0.2 s. After the single presentation of all ten stimulus words with the 30 outcome words, subjects were immediately presented with a test block.

The test blocks were identical to the learning blocks with one exception: instead of typing the letters "A", "B", or "C" to produce the outcome words on the screen, a stimulus word appeared on the screen along with one of "A", "B", or "C", and participants were required to type in the outcome word corresponding to that stimulus-response pair. Together with feedback on their answer, the correct association was shown to the participants until they pressed "ENTER", when the next stimulus word was presented. Once the test block was completed, participants immediately moved to a second learning block in which the same stimulus words were presented in a different order. Across the four test blocks, possible overall scores ranged from 0 to 120.

2.3.4.2. Paired-associates (PA) learning (Williams & Pearlberg, 2006). In this task, participants were presented with 30 pairs of words. A cue word was presented until the participant pressed ENTER, or until 2.5 s elapsed, after which the cue's pair appeared on the screen. They then remained together on screen, again until the participant pressed ENTER, or until 2.5 s elapsed, after which both disappeared and the next cue word was displayed. The test phase was identical to training, except instead of pressing "ENTER" to view the second word of each pair, subjects were required to type that word. Together with feedback on their answer, the correct association was shown to the participant until they pressed "ENTER", when the next word cue was presented. Once the test phase was completed, participants immediately moved to a second learning block in which the same stimulus words were presented in a different order. In total, there were four learning and four test blocks, with possible overall scores ranging from 0 to 120.

Each participant's explicit associative learning score was calculated by summing the 3-Term and PA learning scores.

Please cite this article as: Kaufman, S.B., et al., General intelligence predicts reasoning ability even for evolutionarily familiar content, *Intelligence* (2011), doi:10.1016/j.intell.2011.05.002

2.3.5. Working memory

2.3.5.1. Operation Span Task (Turner & Engle, 1989). The Operation Span (Ospan) task requires participants to store a series of unrelated words in memory while simultaneously solving a series of simple math operations, such as "Is $(9/3)-1=1$?". After participants selected the answer, they were presented with a word (e.g., DOG) to recall. Then participants moved on to the next operation-word string. This procedure was repeated until the end of a set, which varied from two to six items in length. Participants were then prompted to recall all the words from the past set in the same order in which they were presented by typing each word into a box, and using the up and down arrow keys to cycle through the boxes.

Before the main task, participants encountered three practice problems with set size two, where they received feedback about their performance. During these practice trials, we calculated for each participant how long it took them to solve the math operations. Consistent with the methodology of the Automated Ospan task (Unsworth, Heitz, Schrock, & Engle, 2005), we did this to control for individual differences in the time required to solve the math operations. Their mean performance time to solve the equations, plus 2.5SD was used as the time limit for the presentation of the math equations during the main task.

The Ospan score is the sum of all correctly recalled words in their correct positions. The number of operation word-pairs in a set was varied between two, three, four, five, and six with three sets of each. Overall score could range from 0 to 60. Prior research has demonstrated significant correlations between Operation Span and *g* (e.g., Unsworth & Engle, 2005) and a high loading of Operation Span on a general working memory factor (Kane et al., 2004). Each subject's working memory was calculated by summing the Ospan scores for all set sizes.

2.3.6. Processing speed tests

2.3.6.1. Verbal speed test (Speed-V): an English adaptation of a sub-test from the Berlin model of Intelligence Structure (BIS; Jaeger, 1982, 1984). The task was to fill in the missing letter from a 7-letter word; 60 s were given to complete the 57 items. The score is the number completed correctly in 60 s.

2.3.6.2. Numerical speed test (Speed-N): the Speed of Information Processing sub-test from the British Ability Scales (Elliot, 1996). The task was to cross out the highest number in each row of five numbers; 60 s were given to complete 48 items. The score is the number completed correctly in 60 s.

2.3.6.3. Figural speed test (Speed-F). Digit-Symbol, Coding, a sub-test of the WAIS-R that loads on the "processing speed" factor (Deary, 2001). The test was to enter the appropriate symbol (given by a key at the top of the form) beneath a random series of digits; 90 s were given to complete 93 items. The score is the number completed correctly in 90 s.

Processing speed was calculated for each participant by summing Speed-V, Speed-N, and Speed-F.

2.4. Missing values

Some participants were missing values for certain variables. Wherever possible, we estimated their scores using expectation-maximization based on other relevant scores that were available. Due to computer error, values were missing for 9 participants for one of the three second-stratum ability tests (abstract perceptual reasoning, verbal reasoning, and mental rotation ability). Data from the other two markers of second-stratum abilities were used to impute 7 missing abstract perceptual reasoning values, 1 missing verbal reasoning value, and 1 mental rotation ability value. For Speed-F, 3 participants did not follow the directions correctly and their scores could not be included in the analysis. Therefore, we used data from the other two markers of processing speed (Speed-V and Speed-N) to impute 3 missing values on Speed-F. Due to a computer error, performance on the last trial of PA was not recorded for one participant. Since this participant achieved a maximum score on the third trial, we estimated that performance on the last trial was also a perfect score. Finally, one participant was missing all of their explicit associative learning scores, so their explicit associative learning score could not be estimated.

2.5. Calculation of g

An estimate of *g* was calculated by assessing the common variance across abstract perceptual reasoning, verbal reasoning, mental rotation ability, explicit associative learning, working memory, and processing speed, using Principal Axis Factoring (PAF). The first PAF accounted for 41.80% of the total variance in the six scores. (If *g* was calculated only from the three second-stratum ability markers, the first PAF accounted for 69.90% of the variance; *g* calculated in this manner was correlated at $r=.98$ with the *g* used in our analyses.) Table 1 shows the *g*-loadings of all six scores. Because one participant was missing data that could not be estimated for the explicit associative learning tasks, the *N* for *g* is 111.

3. Results

3.1. Correlations

Table 2 shows the correlations among all the cognitive variables. As can be seen, a positive manifold is evident (negative correlations with the deductive reasoning speed tests indicate that faster responders who also choose the correct response tended to score higher on the other cognitive measures).

Table 1
g-loadings of all six scores.

Score	g-loading
Abstract perceptual reasoning	.83
Mental rotation ability	.73
Verbal reasoning	.65
Explicit associative learning	.35
Processing speed	.34
Working memory	.31

Please cite this article as: Kaufman, S.B., et al., General intelligence predicts reasoning ability even for evolutionarily familiar content, *Intelligence* (2011), doi:10.1016/j.intell.2011.05.002

Table 2
Correlations among all cognitive variables.

	1	2	3	4	5	6	7	8	9	10	11	12
1. Abstract perceptual reasoning	–											
2. Verbal reasoning	.50	–										
3. Mental rotation ability	.66	.48	–									
4. Explicit associative learning	.28	.31	.16	–								
5. Working memory	.20	.29	.19	.10	–							
6. Processing speed	.27	.16	.25	.21	.19	–						
7. Arbitrary reasoning– accuracy	.24	.22	.09	.08	.18	.18	–					
8. Precautionary reasoning– accuracy	.30	.32	.24	.21	.25	.21	.28	–				
9. Social exchange reasoning– accuracy	.36	.33	.33	.26	.28	.26	.42	.63	–			
10. Arbitrary reasoning– speed	.13	.05	.13	−.01	−.03	−.13	−.02	.01	.08	–		
11. Precautionary reasoning– speed	−.04	−.17	−.05	−.18	−.24	−.21	−.21	−.27	−.33	.66	–	
12. Social exchange reasoning– speed	−.03	−.10	−.01	−.10	−.18	−.23	−.19	−.25	−.16	.76	.76	–

Note: Correlations > .18 in absolute value are significant at $p < .05$. All correlations had an N of 112, except for correlations with explicit associative learning which had an N of 111.

3.2. Descriptive statistics

Fig. 1a shows the mean proportion correct for both individual cards and whole scenarios (i.e., all cards correct for each scenario). The latter metric was included to render our results more directly comparable with previous literature on the Wason selection task, in which the 4 possible cards for each scenario are typically presented simultaneously, and a scenario is scored correct only if the correct two cards are selected. Examined in this manner, accuracy in our sample for both arbitrary and non-arbitrary problems was very similar to that in previous studies (Stanovich & West, 1998). Fig. 1b shows mean response time for cards only.

In terms of accuracy, proportion correct on the arbitrary trials was significantly less than proportion correct on both precautionary [$t(111) = 17.4$, $p<01$] and social exchange [$t(111) = 17.1$, $p<.01$] reasoning trials. There was no significant difference between proportion correct on precautionary and social exchange reasoning trials. In terms of speed to arrive at the correct answer, the same pattern emerged: mean reaction time to arrive at the correct answer was significantly higher for arbitrary trials than either precautionary [$t(111) = 6.1$, $p<.01$] or social exchange [$t(111) = 5.3$] trials. There was no significant difference between mean RT for precautionary and social exchange trials. The same pattern was found looking at the correlations among the different reasoning types. For both speed and accuracy, precautionary and social exchange performance were more highly correlated with each other than with the arbitrary items (see Table 3). To examine whether this was related to the different

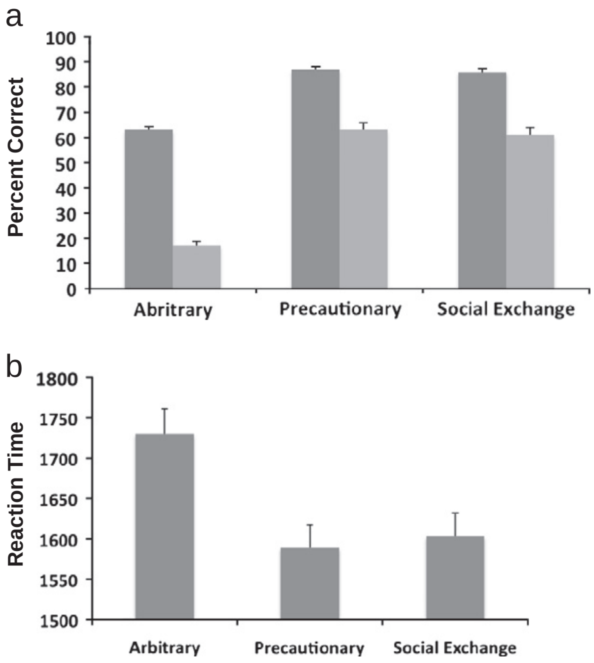


Fig. 1. (a) Mean proportion correct for individual cards (dark grey) and whole scenarios (all 5 cards; light grey), and (b) mean response time (ms) by condition with bars representing S.E. of the mean ($N = 112$).

Please cite this article as: Kaufman, S.B., et al., General intelligence predicts reasoning ability even for evolutionarily familiar content, *Intelligence* (2011), doi:10.1016/j.intell.2011.05.002

Table 3
Correlations among reasoning types for (a) accuracy and (b) speed of correct response. Correlations in parentheses corrected for attenuation.

Measure	Arbitrary	Precautionary	Social exchange
(a)			
Arbitrary	–		
Precautionary	.28** (.49)	–	
Social exchange	.42** (.71)	.63** (.90)	–
α	.48	.67	.73
(b)			
Arbitrary	–		
Precautionary	.66** (.89)	–	
Social exchange	.76** (1.0)	.76** (1.0)	–
Spearman–Brown	.68	.81	.78

reliabilities of the measures for different reasoning types, we also report correlations corrected for attenuation by unreliability (corrected $r = \frac{r_{ab}}{\sqrt{\alpha_a}\sqrt{\alpha_b}}$). Note in Table 3 that all three types of reasoning were more highly correlated with each other when speed is assessed compared to accuracy.

This pattern of results in the sample as a whole, in which performance on both the precautionary and social exchange items is greater than on the arbitrary items for both accuracy and speed is consistent with the results of Reis et al. (2007). Even so, the accuracy rates were lower and the responses were slower than in their sample. This is most likely due to the fact that their participants were older (college students vs. school-age). Nonetheless, the results of both this study and that of Reis et al. (2007) are consistent with prior research showing that at the group level of analysis, precautionary and social exchange reasoning are easier for participants than reasoning that is only arbitrarily contextualized (Evans, 2008).

3.3. Reliability and factor analysis

Looking at the accuracy of responses, the alpha reliability of all 70 trials on the task is .88. Therefore, collapsing across trial type, there is considerable variance that is common among all the trials. In fact, such a high alpha suggests that all of the items, regardless of content (arbitrary, precautionary, or social exchange) tap into a more general contextualized deductive-reasoning accuracy factor. We explored this possibility further by examining a latent contextualized reasoning factor (separately for accuracy and speed) that represents the common variance across the total scores for the three types of reasoning. Using Principal Axis Factoring of the accuracy scores, the first factor (the common variance for mean proportion correct across the three tests) accounted for 63.4% of the total variance. The three loadings on this factor were: arbitrary (.43), precautionary (.65), and social exchange (.97). Note that the lower loading for arbitrary reasoning is due in part to its lower reliability. Table 1 shows that, when corrected for attenuation for unreliability, correlations of precautionary and social exchange with arbitrary become considerably higher. Factor scores were calculated using the regression method. For reaction times, the first factor accounted for 81.7% of the total variance. All three tests loaded extremely highly on this factor: arbitrary (.81), precautionary (.82), and social exchange (.93). Although the correlation between estimated factor scores for the accuracy factor and the speed factor was not significant, there was a trend for higher accuracy to be associated with faster responses ($r = -.17$, $p = .08$), suggesting that there was no speed-accuracy tradeoff.

3.4. Accuracy

Table 4 lists the correlations between accuracy of deductive reasoning and *g*, which was significantly associated with all three forms of deductive reasoning (arbitrary, precautionary, and social exchange), as well as the deductive reasoning accuracy factor. Although *g*'s relation to arbitrary and precautionary reasoning was lower than its relation to social exchange reasoning, this is most likely an artifact of reduced reliability of assessment, given the higher internal consistency for social exchange reasoning (Table 3). It should be noted that the correlations among *g* and accuracy of reasoning across the three types of reasoning were little affected by controlling for processing speed.

In order to determine the *g*-loading of the deductive reasoning tests, we added the three deductive reasoning accuracy items to a principal axis factoring of all the cognitive variables. Social exchange reasoning loaded highly (.69) on *g*, precautionary reasoning loaded moderately (.58), and arbitrary reasoning had the lowest loading (.39) on *g*, but was still substantial. Again, the different loadings are likely to reflect, at least in part, the differing reliability of the deductive reasoning tasks.

3.5. Speed

Table 4 lists the correlations between speed of deductive reasoning and *g*. Here, a different pattern emerged than what was found for accuracy; *g* was not related to any measure of speed of deductive reasoning. Perhaps unsurprisingly, Table 2 suggests that, of our measures of cognitive ability, processing speed demonstrated the most extensive correlations with speed of deductive reasoning, correlating significantly with precautionary and social exchange reasoning.

4. Discussion

The results of the current study suggest that, when individual differences are assessed using psychometrically sound methods that involve aggregating across a sufficiently large number of items, performance on contextualized deductive reasoning problems shows reliable and consistent individual differences, regardless of whether the problems involve arbitrary or evolutionarily relevant rules. Further, accuracy of contextualized deductive reasoning, across

Table 4
Correlations between accuracy and speed of deductive reasoning and *g* ($N = 111$).

Measure	Arbitrary	Precautionary	Social exchange	Factor
Accuracy				
g	.25**	.36**	.44**	.45**
Speed				
g	.12	−.12	−.07	−.04

Notes: *$p<.05$, **$p<.01$.

Please cite this article as: Kaufman, S.B., et al., General intelligence predicts reasoning ability even for evolutionarily familiar content, *Intelligence* (2011), doi:10.1016/j.intell.2011.05.002

content types, is significantly associated with general intelligence. These results support our hypothesis that domain general cognitive abilities should facilitate the solution of any explicit cognitively complex problem, regardless of whether it is additionally facilitated by evolved modular heuristics. The results of this study directly contradict Kanazawa's (2004, 2010) hypothesis that intelligence should be related to performance only on "evolutionarily novel" problems.

In addition to accuracy, we also examined speed and found that *g* is not associated with speed of correct reasoning. Processing speed, however, is associated with speed of correct reasoning, as might be expected since general processing speed should facilitate speed on any particular task. This finding suggests the possibility that domain-general cognitive mechanisms other than *g* can facilitate the speed of deductive reasoning. This is consistent with the study of Reis et al. (2007), who found that the ability to make judgments about emotions predicted speed of reasoning. Further research should investigate the differing cognitive mechanisms that underlie accuracy vs. speed of deductive reasoning.

Our results have implications for several areas of interest to psychologists. Here we discuss their implications for evolutionary psychology, intelligence, and reasoning.

4.1. Evolutionary psychology

As noted in the Introduction, there is an apparent tension between theories of general intelligence and evolutionary psychology's focus on modularity. The results of the current study suggest that both perspectives can be accommodated. On average, in the sample as a whole, participants indeed found the precautionary and social exchange items easier to solve than the arbitrary items, supporting the existence of evolved mechanisms to support reasoning about specific classes of fitness-related problems. At the level of individual differences, however, performance on all three types of content were strongly related to each other, producing a high internal consistency (Cronbach's alpha) for all the items taken together, and performance was significantly related to *g*, supporting the importance of *g* in explaining individual differences in reasoning. Although the loadings of arbitrary and precautionary reasoning on the latent accuracy factor were lower than that for social exchange reasoning, this was most likely due to the lower reliability of the arbitrary and precautionary measures.

As noted above, Kanazawa has hypothesized that what is called "general intelligence" evolved to solve only evolutionarily novel problems. Although a theory of the evolution of a particular mechanism does not translate easily into hypotheses about individual differences (Borsboom & Dolan, 2006), the logic behind Kanazawa's hypothesis that individual differences in *g* should be associated only with performance on evolutionarily novel problems seems to rely on the premise that individual differences in an evolved cognitive ability will be reflected in performance only on the type of problem that the ability evolved to solve. This premise overlooks the existence of exaptation, in which traits evolved for one purpose are eventually used for other purposes (Andrews, Gangestad, & Matthews, 2002; Gould, 1991). In any case, the results of the current study suggest that a stark contrast between "evolutionarily novel" and "evolutionarily familiar" problems may be misguided when considering individual differences, because *g* was significantly associated with forms of reasoning (precautionary and social exchange) which are considered evolutionarily familiar (Cosmides & Tooby, 2004). On a more speculative note, the correlation between *g* and domains of social reasoning for which there may be evolved modular heuristics is consistent with the theory that general intelligence may have increased in humans due to pressure from increasing complexity in social interactions (e.g., Dunbar & Shultz, 2007; Humphrey, 2003).

According to Kanazawa, "[m]ore intelligent individuals are not better than less intelligent individuals at solving evolutionarily familiar problems, such as those in the domains of mating, parenting, interpersonal relationships, and wayfinding (Kanazawa, 2010, p. 35)." However, our results suggest otherwise, as does research on emotional intelligence (when it is assessed properly as an ability), which indicates that the ability to identify and utilize emotional information effectively is typically associated with general intelligence (Mayer, Salovey, & Caruso, 2004; Roberts, Schulze, & MacCann, 2008; Schulte, Ree, & Carretta, 2004). That Kanazawa is wrong is in fact unsurprising, in light of a century of research on intelligence, because the positive manifold has consistently been found to extend to all tests of explicit reasoning (note however that the positive manifold may break down at implicit cognition, see Kaufman, 2011; Kaufman et al., 2010). This suggests that domain general cognitive mechanisms underlying *g* may be actively involved in *any form* of explicit reasoning, even if more specific psychological mechanisms are also brought to bear on the task at hand.

4.2. Intelligence

The results of the current study also relate to issues in the intelligence field. Some intelligence researchers (e.g., Ceci, 1996; Sternberg, 1997) argue for the importance of context and content in understanding individual differences in intellectual performance. To support their argument, they cite studies showing that individuals with average or even low IQ can still reason complexly on "practical" problems. For instance, Cianciolo et al. argue that "practical intelligence," is "an ability—distinct from general or academic intelligence—to perform successfully in naturalistic settings in a way that is consistent with one's goals (Cianciolo et al., 2006; p. 236)." However, it is likely that practical intelligence is at least partially subserved by the same mechanisms as other forms of intelligence (see Gottfredson, 2002), because performance on all problems, regardless of content, tend to be correlated with one another and with *g*.

A major limitation of the practical intelligence studies is that "task complexity" is poorly defined and not controlled or matched to the complexity of IQ tests. The Wason selection task used in the current study can overcome this limitation since the underlying logic system remains constant while task content (e.g., arbitrary vs. non-arbitrary) varies across problems. The distinction between our arbitrary reasoning problems and the non-arbitrary problems seems to correspond to the distinction made in theories of practical intelligence between problems that are unrelated to daily life and those that are. The results of the current study suggest that *g* and the domain general cognitive mechanisms

Please cite this article as: Kaufman, S.B., et al., General intelligence predicts reasoning ability even for evolutionarily familiar content, *Intelligence* (2011), doi:10.1016/j.intell.2011.05.002

associated with *g* may play a larger role in contextualized forms of reasoning than has previously been thought.

The findings of the current study are actually quite consistent with the results of Cianciolo et al. (2006), who assessed the structural relation of *g* to a latent construct they labeled "Practical Intelligence," consisting of the common variance across three everyday tacit-knowledge inventories and the quantitative, verbal, and figural content composites of the Practical subscale of the Sternberg Triarchic Abiltities Test (STAT; Sternberg & T. R. P. C., 2006). They found that all of the indicators of practical intelligence loaded substantially on the Practical Intelligence latent factor. Further, they found a correlation of .48 between their latent Practical Intelligence factor and *g*. They concluded that, "The high-moderate correlation between Practical Intelligence and *g* reflects common variance that may be due to shared demand for neurological functioning and/or shared performance requirements (i.e., test-taking vs. other types of performance)" (p.249). Further, they describe their "Practical Intelligence" latent variable as representing a "general ability to learn from everyday experience (p. 237)."

The results of the current study are consistent with the idea that practical reasoning (precautionary and social exchange) is easier than arbitrary reasoning; participants solved the more practical social exchange and precautionary problems at a much higher rate than the arbitrary problems. Nonetheless, all three types of problem were strongly *g*-loaded. Therefore, intelligence researchers who argue for the separability of practical intelligence from analytical intelligence (e.g., Sternberg, 1997) may be ignoring processes common to both forms of explicit reasoning, whereas intelligence researchers who focus solely on *g* (e.g., Jensen, 1998) may be ignoring the importance of evolutionarily evolved biases and domain-specific mechanisms unique to each form of reasoning. The methodology and findings of the current paper suggests that the two approaches within the intelligence field can be reconciled with one another.

4.3. Reasoning

Lastly, the findings of the current study relate to issues in the reasoning and rationality literature. A number of studies have shown that humans often deviate from normative responses on many reasoning tasks and individual differences in *g* have been associated with the ability to find normatively correct solutions across a range of decision making tasks (Stanovich & West, 2000; Stanovich, 1999; Stanovich & West, 2003; Kokis, Macpherson, Toplak, West, & Stanovich, 2002). Some researchers have argued that this is just further evidence for the positive manifold (consistent positive correlations) found across diverse measures of abstract cognitive ability (e.g., Hunt, 2000), whereas other researchers (e.g., Stanovich & West, 1998) have argued that *g* will play the strongest role on abstract or decontexualized forms of reasoning.

It is important to note however that from a methodological standpoint, most of the prior work on these issues have been on the importance of decontextualized reasoning styles that foster "the tendency to evaluate argument and evidence in a way that is not contaminated by one's prior beliefs" (Stanovich & West, 1997, p. 342). By finding that *g* was significantly associated with arbitrary items (which presumably should not be contaminated by one's prior beliefs) *as well as* non-arbitrary items, the results of the current study are consistent with the results of Klaczynski (2001) and Dominowski and Dallob (1995) and suggest that important individual differences in non-arbitrary contextualized deductive reasoning can be found when a reliable measure of contextualized deductive reasoning is employed. Future research should investigate the role of *g* using reliable measures of contextualized reasoning, employing a wider range of contextualized forms of reasoning than what was used in the current study to investigate the full reach of the positive manifold.

Nonetheless, the finding in the current study that systematic individual differences were related to performance in reasoning do support Stanovich and West's (2000) claim that the positive correlation between *g* and reasoning suggests that the normative model is being applied to evaluate performance and that those with fewer algorithmic limitations (i.e., higher *g*) can be predicted to come closer to the "rational response." Given the data in the current study, one cannot argue that all errors were simply the result of random lapses in attention or that all participants were incapable of being rational. Clearly, some participants were better at accurately reasoning than others, and individual differences in *g* were important in distinguishing performance.

4.4. Limitations

The current study does have some limitations that should be overcome in future studies. First, the version of the Wason card task that was administered in the current study was time limited. Having to solve the problem in 4 s may have strengthened reliance on *g* mechanisms. However, such a short amount of time probably minimized explicit reasoning, and controlling for processing speed did not affect correlations with *g*. Additionally, in the sample as a whole, participants found the precautionary and social exchange reasoning problems easier to solve even though they were also the more *highly* *g*-loaded tasks (compared to the arbitrary items). Further, level of accuracy on the precautionary and social exchange items was above 85%. Such a high level of accuracy suggests that even with more time allowed to solve the problems, there isn't much more variance in task performance that could come about. Finally, the functioning of cognitive modules that are evolved to process specific types of information automatically should be faster than the effortful cognitive processes associated with *g*. Nevertheless, future research should alter the time limitations of the task to see if correlations with *g* are affected.

Another limitation of the current study is that purely symbolic, abstract deductive reasoning items weren't administered. Because the main purpose of our study was to assess the role of *g* in different forms of contextualized deductive reasoning, we kept the contextualized format of presentation constant and varied only the *kind* of reasoning. As a result, our study cannot directly compare the strength of the correlation between *g* and abstract vs. contextualized forms of reasoning (à la Stanovich & West, 1998). Future research should include abstract reasoning items as well so that the full reach of the positive manifold can be assessed on an even broader set of deductive reasoning items.

Please cite this article as: Kaufman, S.B., et al., General intelligence predicts reasoning ability even for evolutionarily familiar content, *Intelligence* (2011), doi:10.1016/j.intell.2011.05.002

A third limitation is the higher than average intelligence of the current sample. Future research should look at samples with a wider-range of cognitive ability. With that said, the findings in the current study are all the more impressive considering the restricted sample. A less restricted sample may produce even higher *g*-loadings for contextualized deductive reasoning tasks.

5. Conclusion

Despite the fact that people are better at reasoning about non-arbitrary rules in evolutionarily relevant situations than they are at reasoning about arbitrary rules, performance on arbitrary and non-arbitrary problems are related and reflect an underlying latent variable of reasoning ability. This ability is related to general intelligence. Contrary to a hypothesis by Kanazawa (2010), performance on non-arbitrary, evolutionarily familiar problems is more strongly related to *g* than performance on arbitrary, evolutionary novel problems. These results should prompt evolutionary psychologists to rethink the importance of general intelligence in human mental processes, while affirming that human beings may have evolved cognitive modules that facilitate reasoning about specific, evolutionarily relevant situations. In addition to those specific modules, people may draw on domain general cognitive mechanisms when solving any reasonably complex cognitive problem. This would be consistent with the theory that, as a species-typical trait, general intelligence may have evolved in response to all complex situations in which modular heuristics did not reliably generate the optimal response.

Acknowledgements

The authors would like to thank Nicholas J. Mackintosh for his feedback and assistance with the study design and analysis, Sheila Bennett for her kind assistance in recruiting participants, Jim Blair and Nikhil Srivastava for computer support, and the administration at Hills Road Sixth Form College for the use of their facilities.

References

Andrews, P. W., Gangestad, S. W., & Matthews, D. (2002). Adaptionism – how to carry out an exaptationist program. *Behavioral and Brain Sciences*, *25*, 489–504.

Arden, R., Gottfredson, L. S., Miller, G., & Pierce, A. (2008). Intelligence and semen quality are positively correlated. *Intelligence*, *37*, 277–282.

Barrett, H. C. (2009). Where there is an adaptation, there is a domain: The form-function fit in information processing. In S. M. Platek, & T. K. Shackelford (Eds.), *Foundations in evolutionary cognitive neuroscience* (pp. 97–116). Cambridge, UK: Cambridge University Press.

Bartholomew, D. J., Deary, I. J., & Lawn, M. (2009). A new lease of life for Thomson's bonds model of intelligence. *Psychological Review*, *116*, 567–579.

Borsboom, D., & Dolan, C. V. (2006). Why g is not an adaptation: A comment on Kanazawa (2004). *Psychological Review*, *113*, 433–437.

Buss, D. M., & Hawley, P. H. (Eds.). (2010). *The evolution of personality and individual differences*. New York, NY: Oxford University Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, UK: Cambridge University Press.

Ceci, S. J. (1996). *On intelligence: A bioecological treatise on intellectual development.* Boston, MA: Harvard University Press.

Chabris, C. F. (2007). Cognitive and neurobiological mechanisms of the Law of General Intelligence. In M. J. Roberts (Ed.), *Integrating the mind: Domain general vs domain specific processes in higher cognition* (pp. 449–491). New York, NY: Psychology Press.

Cheng, P., & Holyoak, K. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.

Chiappe, D., & MacDonald, K. (2005). The evolution of domain-general mechanisms in intelligence and learning. *The Journal of General Psychology*, *132*, 5–40.

Cianciolo, A. T., Grigorenko, E. L., Jarvin, L., Guillermo, G., Drebot, M. E., & Sternberg, R. J. (2006). Practical intelligence and tacit knowledge: Advancements in the measurement of developing expertise. *Learning and Individual Differences*, *16*, 235–253.

Corporation, T. P. (1995). *DAT for selection-verbal reasoning (European adaptation).* Orlando, FL: Harcourt Brace and Company.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, *31*, 187–276.

Cosmides, L., Barrett, H. C., & Tooby, J. (2010). Adaptive specializations, social exchange, and the evolution of human intelligence. *Proceedings of the National Academy of Sciences*, *107*, 9007–9014.

Cosmides, L., & Tooby, J. (2001). Unraveling the enigma of human intelligence: Evolutionary psychology and the multimodular mind. In R. J. Sternberg, & J. C. Kaufman (Eds.), *The evolution of intelligence* (pp. 145–198). Hillsdale, NJ: Erlbaum.

Cosmides, L., & Tooby, J. (2004). Social exchange: The evolutionary design of a neurocognitive system. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1295–1308). (3 rd ed.). Cambridge, MA: MIT Press.

Deaner, R. O., Isler, K., Burkart, J., & van Schaik, C. (2007). Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates. *Brain, Behavior and Evolution*, *70*, 115–124.

Deaner, R. O., van Schaik, C. P., & Johnson, V. (2006). Do some taxa have better domain- general cognition than others? A meta-analysis. *Evolutionary Psychology*, *4*, 149–196.

Deary, I. (2001). Human intelligence differences: A recent history. *Trends in Cognitive Sciences*, *5*, 127–130.

Dominowski, R. L., & Dallob, P. (1995). Insight and problem solving. In R. J. Sternberg, & J. E. Davidson (Eds.), *The nature of insight* (pp. 33–62). Cambridge, MA: The MIT Press.

Dunbar, R. I. M., & Shultz, S. (2007). Understanding primate brain evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*, 649–658.

Elliot, C. D. (1996). *British Ability Scales II.* Windsor, Berkshire, UK: NFER-Nelson.

Ermer, E., Guerin, S. A., Cosmides, L., Tooby, J., & Miller, M. B. (2006). Theory of mind broad and narrow: Reasoning about social exchange engages ToM areas, precautionary reasoning does not. *Social Neuroscience*, *1*, 196–219.

Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*, 454–459.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278.

Fiddick, L., Cosmides, L., & Tooby, J. (2000). No interpretation without representation: The role of domain-specific representations and inferences in the Wason selection task. *Cognition*, *77*, 1–79.

Fiddick, L., Spampinato, M. V., & Grafman, J. (2005). Social contracts and precautions activate different neurological systems: An fMRI investigation of deontic reasoning. *NeuroImage*, *28*, 778–786.

Furlow, F. B., Armijo-Prewitt, T., Gangestad, S. W., & Thornhill, R. (1997). Fluctuating asymmetry and psychometric intelligence. *Proceedings of the Society of London, Series B*, *264*, 823–829.

Geary, D. C. (2004). *Origin of mind: Evolution of brain, cognition, and general intelligence.* Washington, D.C.: American Psychological Association.

Geary, D. C. (2009). Evolution of general fluid intelligence. In S. M. Platek, & T. K. Shackelford (Eds.), *Foundations in evolutionary cognitive neuroscience* (pp. 25–26). Cambridge, MA: MIT Press.

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition*, *43*, 127–171.

Girotto, V., & Tentori, K. (2008). Is domain-general thinking a domain-specific adaptation? *Mind & Society*, *7*, 167–175.

Gottfredson, L. S. (2002). g: Highly general and highly practical. In R. J. Sternberg, & E. L. Grigorenko (Eds.), *The general factor of intelligence: How general is it?* (pp. 331–380). Mahwah, NJ: Lawrence Erlbaum.

Gould, S. J. (1991). Exaptation: A crucial tool for evolutionary psychology. *Journal of Social Issues*, *47*, 43–65.

Haig, B. D. (2005). Exploratory factor analysis, theory generation, and scientific method. *Multivariate Behavioral Research*, *40*, 303–329.

Humphrey, N. (2003). *The inner eye: Social intelligence in evolution.* New York, NY: Oxford University Press.

Hunt, E. (2000). Situational constraints on normative reasoning. *The Behavioral and Brain Sciences*, *23*, 680.

Jaeger, A. O. (1982). Mehrmodale Klassifikation von Intelligenzleistungen: experimentell kontrollierte Weiterentwicklungeines deskriptiven Intelligenzstructurmodells. [Multimodal classification of cognitive abilities:

an experimentally controlled development of a descriptive model of intelligence structure]. *Diagnostica*, *28*, 195–225.

Jaeger, A. O. (1984). Intelligenzstructurforschung: Konkurriende Modelle, neue Entwicklungen, Perspectiven [Research in the structure of intelligence: competing models, new developments, perspectives]. *Psychologische Rundschau*, *35*, 21–35.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

Johnson, W., & Bouchard, T. J., Jr. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393–416.

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, *30*, 135–154.

Kanazawa, S. (2004). General intelligence as a domain-specific adaptation. *Psychological Review*, *111*, 512–523.

Kanazawa, S. (2010). Evolutionary psychology and intelligence research. *American Psychologist*, *65*, 279–289.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*, 189–217.

Kaufman, S. B. (2011). Intelligence and the cognitive unconscious. In R. J. Sternberg, & S. B. Kaufman (Eds.), *The Cambridge Handbook of Intelligence*. New York, NY: Cambridge University Press.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Brown, J., & Mackintosh, N. (2009). Associative learning predicts intelligence above and beyond working memory and processing speed. *Intelligence*, *37*, 374–382.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jimenez, L., Brown, J. B., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*, 321–340.

Keller, M., & Miller, G. F. (2006). Which evolutionary genetic models best explain the persistence of common, harmful, heritable mental disorders? *The Behavioral and Brain Sciences*, *29*, 385–404.

Klaczynski, P. A. (2001). Analytic and heuristic processing influences on adolescent reasoning and decision-making. *Child Development*, *72*, 844–861.

Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology*, *83*, 26–52.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry*, *60*, 197–215.

Miller, G. F. (2000). How to keep our metatheories adaptive: Beyond Cosmides, Tooby, and Lakatos. *Psychological Inquiry*, *11*, 42–46.

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, *61*, 622–631.

Penke, L. (2010). Bridging the gap between modern evolutionary psychology and the study of individual differences. In D. M. Buss, & P. H. Hawley (Eds.), *The evolution of personality and individual differences* (pp. 243–279). New York, NY: Oxford University Press.

Penke, L., Borsboom, D., Johnson, W., Kievit, R. A., Ploeger, A., & Wicherts, J. M. (in press). Evolutionary psychology and intelligence research cannot be integrated the way Kanazawa (2010) suggests. American Psychologist.

Peterson, J. B. (1999). *Maps of meaning: The architecture of belief.* New York, NY: Routledge.

Peterson, J. B., & Flanders, J. L. (2002). Complexity management theory: Motivation for ideological rigidity and social conflict. *Cortex*, *38*, 429–458.

Prokosch, M. D., Yeo, R. A., & Miller, G. F. (2005). Intelligence tests with higher *g*-loadings show higher correlations with body symmetry: Evidence for a general fitness factor mediated by developmental stability. *Intelligence*, *33*, 203–213.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales.* Oxford, UK: Oxford Psychologists Press.

Reis, D. L., Brackett, M. A., Shamosh, N. A., Kiehl, K. A., Salovey, P., & Gray, J. R. (2007). Emotional intelligence predicts individual differences in social exchange reasoning. *NeuroImage*, *35*, 1385–1391.

Roberts, R. D., Schulze, R., & MacCann, C. (2008). The measurement of emotional intelligence: A decade of progress? In G. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The Sage Handbook of Personality Theory and Assessment, Vol. 2.* (pp. 461–482)Los Angeles: Sage.

Schulte, M. J., Ree, M. J., & Carretta, T. R. (2004). Emotional intelligence: not much more than g and personality. *Personality and Individual Differences*, *37*, 1059–1068.

Spearman, C. (1904). “General intelligence,” objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In L. A. Hirschfeld, & S. Gelman (Eds.), *Mapping the mind: Domain-specificity in cognition and culture* (pp. 39–67). Cambridge, UK: Cambridge University Press.

Stanovich, K. (1999). *Who is rational?: Studies of individual differences in reasoning.* Mahwah, NJ: Lawrence Erlbaum.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, *89*, 342–357.

Stanovich, K. E., & West, R. F. (1998). Cognitive ability and variation in selection task performance. *Thinking & Reasoning*, *4*, 193–230.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *The Behavioral and Brain Sciences*, *23*, 645–726.

Stanovich, K. E., & West, R. F. (2003). The rationality debate as a progressive research program. *The Behavioral and Brain Sciences*, *26*(4), 531–534.

Sternberg, R. J. (1997). *Successful intelligence: How practical and creative intelligence determine success in life.* New York, NY: Plume.

Sternberg, R. J., & T. R. P. C. (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical and creative skills. *Intelligence*, *34*, 321–350.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154.

Unsworth, N., & Engle, R. W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory & Cognition*, *33*, 213–220.

Unsworth, N., Heitz, R., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*, 498–505.

van der Maas, H. L. J., et al. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*, 842–861.

Vandenberg, S. G., & Kruse, A. R. (1978). Mental rotations: Group tests of three-dimensional spatial visualization. *Perceptual and Motor Skills*, *47*, 599–604.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, *3*, 255–274.

Wason, P. C. (1968). Reasoning about a rule. *The Quarterly Journal of Experimental Psychology*, *20*, 273–281.

Williams, B. A., & Pearlberg, S. L. (2006). Learning of three-term contingencies correlates with Raven scores, but not with measures of cognitive processing. *Intelligence*, *34*, 177–191.

Woodley, M. A. (2010). Are high-IQ individuals deficient in common sense? A critical examination of the ‘clever sillies’ hypothesis. *Intelligence*, *38*, 471–480.

Yeo, R. A., Gangestad, S. W., Liu, J., Calhoun, V. D., & Hutchison, K. E. (2011). Rare copy number deletions predict individual variation in intelligence. *PLoS ONE*, *6*, e16339.

Please cite this article as: Kaufman, S.B., et al., General intelligence predicts reasoning ability even for evolutionarily familiar content, *Intelligence* (2011), doi:10.1016/j.intell.2011.05.002