Assessing and Explaining Individual Differences
within the Adaptive Toolbox Framework:

New Methodological and Empirical Approaches
to the Recognition Heuristic

Martha Michalkiewicz

Dipl. Math.

Inaugural dissertation

submitted in partial fulfillment of the requirements for the degree Doctor of Social Sciences

in the Graduate School of Economic and Social Sciences

at the University of Mannheim

Contents

Abstract

Individuals do not only show large differences with regard to the judgments and decisions they make, but also with regard to the strategies they use to arrive at their decisions. However, individual differences in decision strategy selection have gained insufficient attention so far. For this reason, I investigate individual differences with respect to the application of the fast-and-frugal heuristics of the adaptive toolbox – a framework that has become increasingly important within the field of decision making. In particular, I address one of the most prominent examples of the adaptive toolbox: the recognition heuristic (RH), that is, a decision strategy for paired comparisons which bases choice solely on recognition while ignoring any additional information.

The overarching aim of my thesis is to enhance the understanding of the cognitive and personality traits underlying individual differences in use of the RH. However, so far, there has been a deficiency in the methods relating individual traits to RH-use. For this purpose, I extend a measurement model of the RH to a hierarchical version incorporating individual traits directly into the estimation of RH-use. This methodological advance allows detection of the dispositional determinants of variation in strategy selection regarding the RH in a straightforward and unbiased way.

Equipped with the required methods, the first project reported in this thesis investigates temporal and cross-situational stability in use of the RH. By demonstrating these important preconditions, I ensure that it is principally possible to find reliable relations between individual traits and RH-use. Building upon these results, the second project addresses the effect of (fluid and crystallized) intelligence on individual differences in adaptive RH-use. In sum, there is supportive evidence that adaptive application of the RH to the decision context is moderated by fluid but not crystallized intelligence. Extending this line of research, the third project aims at explaining individual differences in RH-use free of any interaction with the situation. In brief, RH-use is found to decrease with need for cognition (i.e., inclination towards cognitively demanding activities) but not to increase with faith in intuition (i.e., trust in feelings).

To conclude, by means of the three projects reported herein and with the aid of the newly developed hierarchical measurement model of RH-use, I demonstrate that RH-use represents a person-specific decision making style that is temporally and cross-situationally stable, and that is affected by fluid intelligence and need for cognition.

1. Manuscripts

This thesis is based on the three manuscripts listed below. In the following sections, I give a short introduction to the theory and especially to the newly developed methods, briefly summarize each of the manuscripts, and provide a general discussion as well as an outlook for future directions. The three manuscripts are attached at the end of this thesis in the same order as presented here.

Manuscript 1

Michalkiewicz, M., & Erdfelder, E. (2016). Individual differences in use of the recognition heuristic are stable across time, choice objects, domains, and presentation formats. *Memory & Cognition, 44,* 454-468. doi:10.3758/s13421-015-0567-6

Manuscript 2

Michalkiewicz, M., Arden, K., & Erdfelder, E. (2016). *Do smarter people make better decisions? The influence of intelligence on adaptive use of the recognition heuristic.* Manuscript submitted for publication.

Manuscript 3

Michalkiewicz, M., Minich, B., & Erdfelder, E. (2016). *Explaining individual differences in fast-and-frugal decision making: The impact of need for cognition and faith in intuition on use of the recognition heuristic.* Manuscript submitted for publication.

2. Introduction

"When optimal solutions are out of reach, we are not paralyzed to inaction or doomed to failure. We can use heuristics to discover good solutions." (Gigerenzer, 2004, p. 63)

We all have to make many decisions and judgments every day. Which orange juice should I buy? Which stocks should I invest in? Which football team should I bet on in the next game? If we do not know the answer, we have to infer it from our knowledge or based on information that we can acquire from the decision environment. In fact, the answers to many questions are often not directly accessible, thus rendering our decisions uncertain. Furthermore, due to a lack of time, limited knowledge and limited cognitive resources, we are forced to make decisions and judgments without constantly retrieving all available knowledge, searching for all existing information in the environment, and consciously evaluating every alternative. But how do we make a decision or a judgment under these circumstances?

Addressing this important question, Gigerenzer, Todd, and the ABC Research Group (1999) introduced the metaphor of the adaptive toolbox. According to this framework, decision makers are equipped with a repertoire of strategies – termed heuristics – to solve the decision problems they face. Heuristics are characterized as domain-specific, meaning that each heuristic is tailored to a specific decision task and situation. Also, heuristics are considered fast-and-frugal, as they require only a minimum of information, time, and cognitive resources (Gigerenzer & Selten, 2001). As such, information search focuses on only a small part of the available information and is terminated by a simple stopping rule. Additionally, choice follows a simple and clearly defined decision rule based on a small amount of information found.

Within the last decades, the adaptive toolbox framework has inspired much innovative research within the field of decision making, and has proposed a large number of heuristics (for an overview, see for instance Gigerenzer & Brighton, 2009; Gigerenzer & Gaissmaier, 2011; Goldstein & Gigerenzer, 2009; Hertwig & Pachur, 2015). In my thesis, I address the recognition heuristic (RH; Goldstein & Gigerenzer, 1999, 2002), which is introduced in the next section. I decided to pay attention to this heuristic for two main reasons. First, the RH has already been intensively studied, which assures a fairly comprehensive theory as a starting point for my research.

Second, the RH represents an essential building block of multi-alternative decision strategies (e.g., Frosch, Beaman, & McCloy, 2007; Marewski, Gaissmaier, Schooler, Goldstein, & Gigerenzer, 2010; McCloy, Beaman, & Smith, 2008), of other fast-and-frugal heuristics (e.g., Gigerenzer & Goldstein, 1999), and of strategies with respect to preferences (Hoyer & Brown, 1990; Jacoby, Kelley, Brown, & Jasechko, 1989; Macdonald & Sharp, 2000; Oeusoonthornwattana & Shanks, 2010). Thus, it seems reasonable to assume that the results of my thesis do not only offer insights into properties of the RH but also broaden our understanding of many other decision strategies.

Note that I do not reiterate all findings on the RH in what follows (for a review, see for instance Gigerenzer & Goldstein, 2011; Pachur, Todd, Gigerenzer, Schooler, & Goldstein, 2011; Pohl, 2011). Instead, I focus on specific aspects that are important for the development of the research questions posed in my thesis.

2.1. The Recognition Heuristic (RH)

"If one of two objects is recognized and the other is not, then infer that the recognized object has the higher value with respect to the criterion." (Goldstein & Gigerenzer, 2002, p. 76)

The RH is one of the simplest and yet most widely investigated strategies of the adaptive toolbox. When deciding between recognized and unrecognized choice objects, the RH exploits a core human ability, namely, recognition memory: Decisions are exclusively determined by whether or not the objects are recognized; any additional knowledge is ignored. For example, when asked to infer which of two cities is more populous, "Tokyo" or "Harbin", according to the RH, a decision maker should choose Tokyo simply because he or she recognizes Tokyo but has not heard of Harbin. The alternative strategy is the integration of knowledge about the recognized object retrieved from memory. To illustrate, a decision maker might choose Tokyo based on her or his knowledge that Tokyo is the Japanese capital and that national capitals are most often more populous than ordinary cities. Both strategies result in the same decision but differ in the underlying cognitive mechanisms.

Individual differences in use of the RH

In general, the RH represents a useful decision strategy as it exploits the systematic link between recognition and the decision criterion inherent in a large number of decision situations. The strength of this link, called recognition validity, determines the accuracy or success of the RH. To be precise, the recognition validity has been defined by Goldstein and Gigerenzer (2002) as the proportion of cases where the recognized object has a higher value than the unrecognized object with respect to the decision criterion (based on all cases where one object is recognized and the other is not). To illustrate, when asked to infer which of two cities is more populous, people should be more likely to recognize more populous cities because they appear more often in the media and in everyday conversations.

For this reason, the RH has been shown to be successful in many decision tasks, often leading to the same amount or even more accurate inferences than the use of more advanced knowledge-based strategies. One important requirement for the RH to be useful, and in fact used, is that the recognition validity is larger than the knowledge validity. Knowledge validity is defined as the proportion of cases where use of knowledge leads to a correct response, or as stated by Goldstein and Gigerenzer (2002), as the probability of a correct response when both objects are recognized. Tasks with this feature include geographical questions like the size of cities or the length of rivers (e.g., Goldstein & Gigerenzer, 2002; Pohl, 2006), as well as economic and political questions like the safety of airlines and the success of politicians (e.g., Gaissmaier & Marewski, 2011; Richter & Späth, 2006). Furthermore, the RH has been found to outperform experts when forecasting sport events (e.g., Ayton, Önkal, & McReynolds, 2011; Pachur & Biele, 2007; Scheibehenne & Bröder, 2007; Serwe & Frings, 2006) and to outperform sophisticated marked indices when predicting stock-market performance (e.g., Borges, Goldstein, Ortmann, & Gigerenzer, 1999; Erdfelder, Küpper-Tetzel, & Mattern, 2011; Newell & Shanks, 2004).[1]

So far, the vast majority of research has focused on specific properties of the decision situation which might affect RH-use. In fact, it is quite well established that decision makers apply the RH in accordance with the decision context. A summary of the corresponding findings is given in the following section.

---

[1] Of course, there is also research showing that the success of the RH can easily disappear under certain conditions (e.g., McCloy, Beaman, Frosch, & Goddard, 2010; Oppenheimer, 2003; Pachur & Hertwig, 2006; Pachur, Bröder, & Marewski, 2008; Pohl, 2011; Richter & Späth, 2006).

2.2. Adaptive RH-Use

"The question is not whether individuals always rely on a given heuristic, but whether they use heuristics in an adaptive way." (Gigerenzer & Brighton, 2009, p. 134)

There is general agreement that decision makers display a great deal of adaptivity. Accordingly, individuals chose among different decision strategies as a function of their own cognitive capacities on the one hand and the requirements of the situation on the other hand (e.g., Gigerenzer & Selten, 2001; Payne, Bettman, & Johnson, 1988, 1993; Simon, 1956, 1990). Building upon these insights, Gigerenzer et al. (1999) acknowledge that individuals have to select decision strategies in accordance with their own resources, meaning available knowledge and available cognitive capacities. At the same time, they recognize the influence of the environmental structure, for instance, in terms of success rates of different decision strategies or time required to execute them. Emphasizing the adaptive aspect of decision making even more, Gigerenzer et al. (1999) apply the concept of ecologic rationality, that is, the match between the structure of a heuristic and the structure of the respective environment. In this sense, the success of a heuristic can only be judged within the corresponding decision context.

Previous work has consistently demonstrated adaptive use of the RH. However, this research mostly focused on situational influences. In particular, decision makers were shown to use the RH adaptively depending on the size of the recognition and the knowledge validities. More precisely, they applied the RH more often the more valid the recognition cue was (e.g., Castela, Kellen, Erdfelder, & Hilbig, 2014; Hilbig, Erdfelder, & Pohl, 2010; Pachur, Mata, & Schooler, 2009; Pohl, 2006; Scheibehenne & Bröder, 2007). In contrast, decision makers applied the RH less, the easier further knowledge was available and could be integrated (e.g., Bröder & Eichler, 2006; Glöckner & Bröder, 2011, 2014; Hilbig, Michalkiewicz, Castela, Pohl, & Erdfelder, 2015; Hilbig, Pohl, & Bröder, 2009; Newell & Fernandez, 2006; Richter & Späth, 2006). Additionally, decision makers were shown to adapt RH-use dependent on the current constraints of the situation. In particular, they applied the RH more often when they had to make inferences under time pressure (Hilbig, Erdfelder, & Pohl, 2012; Pachur & Hertwig, 2006).

Although there is clear evidence that people use decision strategies adaptively, there is also agreement that there are limitations to human adaptivity and that failures in adaptivity do occur (Payne et al., 1993). However, only few studies investigated adaptive RH-use with respect to cognitive limitations, for instance, RH-use under depleted cognitive resources (Pohl, Erdfelder, Hilbig, Liebke, & Stahlberg, 2013). This gap suggests that one essential aspect has been largely neglected in most studies: individual differences in the application of the RH as well as their dispositional determinants. I will take a closer look at these aspects in the next section.

2.3. Individual Differences in RH-Use

"Strategy differences between people appear to be the rule rather than marginal exceptions." (Pachur et al., 2008, p. 205)

Many studies originally tailored to investigate the use of the RH on the group level also revealed large individual differences in RH-use within groups over and above situational influences. Figure 1 illustrates this insight by showing individual-level analyses of studies on RH-use comprising no situational manipulation. In this case, individual RH-use is assessed via the latent-trait r-model described in detail in section 3.2.2. As can be seen in Figure 1, even under constant contextual conditions, participants differ to a large extent with regard to individual RH-use within each experiment. Similar findings have repeatedly been found throughout the literature (e.g., Gigerenzer & Brighton, 2009; Hilbig & Richter, 2011; Marewski et al., 2010; Newell & Shanks, 2004; Pachur et al., 2008). However, until now, there is very limited work on dispositional determinants that might explain this large variation in RH-use.

So far, only Hilbig (2008) successfully investigated determinants of differences in RH-use on the individual level. In particular, he demonstrated that neuroticism is associated with higher levels of adherence to the RH. Apart from this finding, there is evidence that different groups of people might show preferences for different strategies. For instance, Hilbig and Pohl (2008) found that more knowledgeable people tended to refrain from applying the RH compared to less knowledgeable ones. Exploring a different source of individual differences, Pachur et al. (2009) and Pohl, von Massow,

Individual differences in use of the RH

and Beckmann (2016) investigated RH-use across the lifespan. In particular, Pohl et al. (2016) found that adolescents applied the RH more frequently compared to both preadolescent children and young adults, whereas preadolescent children and young adults showed similar levels of RH-use. Additionally, Pachur et al. (2009) showed that elderly people relied on the RH more often compared to young adults (for a reanalysis, see Horn, Pachur, & Mata, 2015). In fact, although the findings concerning age are limited to groups, they support the idea that RH-use might be determined by person-specific characteristics.

To conclude, the examples described above show that individuals significantly differ from each other with regard to the propensity to rely on the RH even under constant contextual conditions. At the same time, the existing research reveals that there is a lack of research on potential explanations waiting to be filled.
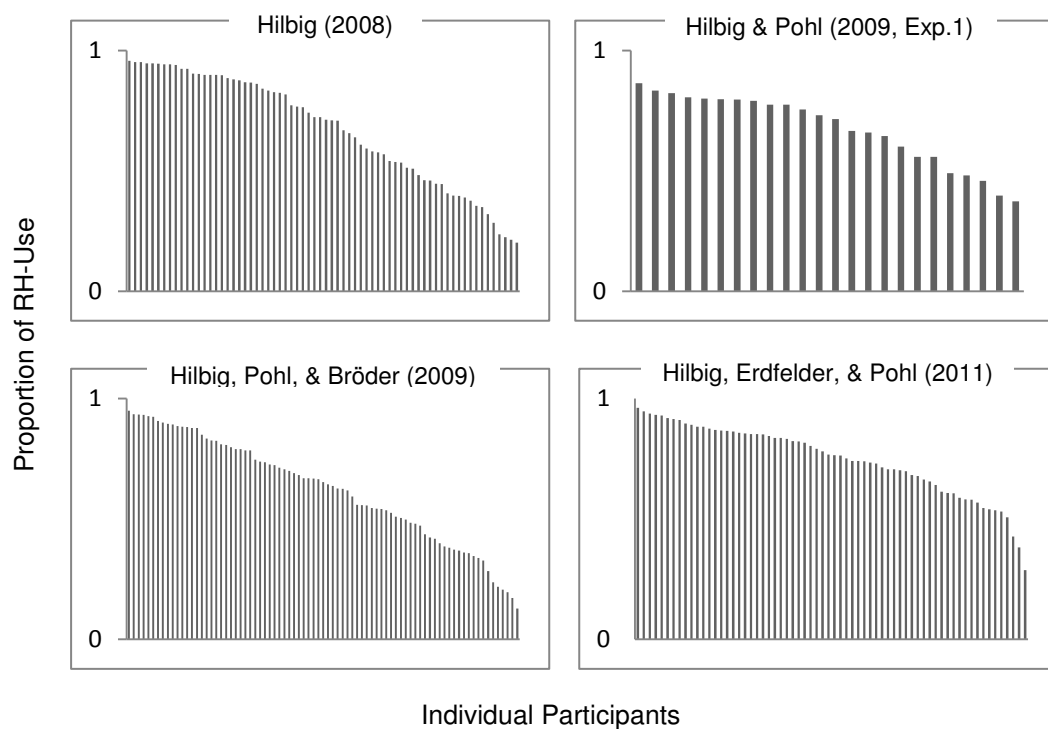
Figure 1. Individual proportions of RH-use per participant for the data from Hilbig (2008), Hilbig and Pohl (2009, Exp.1), Hilbig, Pohl, and Bröder (2009), and Hilbig, Erdfelder, and Pohl (2011). RH-use is estimated by means of the latent-trait r-model (Michalkiewicz & Erdfelder, 2016) and ordered by size.[2]

---

[2] I am grateful to Benjamin Hilbig for providing the raw data of Hilbig (2008), Hilbig and Pohl (2009), Hilbig et al. (2009), and Hilbig et al. (2011).

Individual differences in use of the RH

2.4. Explaining Individual Differences in RH-Use

"Whereas there is substantial evidence that task characteristics influence decisions and initial evidence that individual differences affect various stages of decision making, what is not clear is the nature of relations among situational variables, individual differences, and decision-making processes." (Mohammed & Schwall, 2009, p. 294)

Investigating the effects of individual traits (as well as the interaction of situational variables and individual traits) on decision behavior represents an important goal not only from the perspective of decision making but also from the perspective of personality psychology. To illustrate, there is agreement in the field of decision making that choice of the appropriate decision strategy for a particular problem is largely determined by both the characteristics of the decision problem and the characteristics of the decision maker (e.g., Appelt, Milch, Handgraaf, & Weber, 2011; Mohammed & Schwall, 2009; Payne et al., 1993; Stanovich & West, 2000).[3] Although, for instance, the adaptive decision maker framework (Payne et al., 1993) largely focuses on adaptive behavior with respect to task and context variables, it also makes clear that there are individual differences in decision strategy selection which are moderated by person-specific variables like cognitive ability and prior knowledge. In a similar way, research in personality psychology emphasizes the importance of exploring how personality, situations, and behaviors are interrelated (e.g., Fleeson & Noftle, 2008; Funder, 2001, 2008; Mischel, 1968). In particular, the question is asked whether individual differences in specific personality traits are large enough to have a substantial influence on behavior, and whether this influence persists under varying situational conditions. Hence, the same conclusion is reached from both perspectives: To fully understand human decision behavior it is crucial to understand 1) the situational and 2) the individual determinants of decision strategy selection. However, despite the

---

[3] According to Payne et al. (1993), the social context represents the third major factor of determinants of strategy selection, including concepts like group membership. This factor will not be discussed further as it is of minor interest for the research questions posed in this thesis. For more information on group decision making with respect to heuristic-use, see for instance Reimer and Hoffrage (2003, 2005, 2006), and especially with respect to the RH, see Kämmer, Gaissmaier, Reimer, and Schermuly (2014) and Reimer and Katsikopoulos (2004).

fundamental importance of this research question, little information on the influence of personality (and of the situation-personality interaction) on behavior is available (e.g., Funder, 2001; Mohammed & Schwall, 2009). This lack of information is particularly visible with respect to the RH as summarized in sections 2.2 and 2.3.

Unquestionably, no trait variable in isolation might sufficiently account for the large individual variation in decision strategy selection. Rather, individual differences in decision behavior are evoked by a combination of different traits (e.g., Stanovich & West, 2000). Unquestionably, there are many personality theories that could account for individual differences in decision strategy selection. For instance, Bröder (2012) reported a large series of studies on the influence of personality on use of the take-the-best heuristic (TTB[4]; Gigerenzer & Goldstein, 1999). In the same way, it might be considered to successively investigate these individual traits and even more with respect to the RH. This would result in a comprehensive overview regarding the quest which trait variables determine RH-use and which do not. However, without a strong theory this approach would be based on the try and error principle. Instead, we need to wisely select a promising starting point, that is, a stable trait variable that offers a strong theoretical framework relevant for our decision task and that enables deriving clear and testable predictions (e.g., Mohammed & Schwall, 2009; Phillips, Fletcher, Marks, & Hine, 2016). Specifically, the following two approaches seem most promising.

First, consider the variation in cognitive capacity as an explanation for individual differences in adaptive decision making. In fact, cognitive capacity is often assumed to represent the most fundamental predictor of reasoning (e.g., Stanovich & West, 1998, 2000). The best way to operationalize cognitive capacity is in terms of its most fundamental form, namely, intelligence. Indeed, measures of general intelligence have been shown to be linked to practically all sub-processes belonging to human cognition (e.g., Carroll, 1993; Kane, Conway, & Hambrick, 2005). Therefore, I explore the effect of intelligence on adaptive RH-use, in line with the idea that adaptivity is generally interpreted as a sign of intelligence (e.g., Larrick, Nisbett, & Morgan, 1993;

---

[4] The TTB heuristic sequentially compares the cues of two or more alternatives, ordered by cue validity (from the highest to the lowest), and chooses the alternative that is superior with regard to the first cue that discriminates between alternatives.

Individual differences in use of the RH

Neisser et al., 1996; Payne et al., 1993; Sternberg & Salter, 1982). The results of this investigation are summarized in section 4.2.

Second, consider information processing styles (or dual process models of information processing) as an explanation for individual differences in decision strategy selection. According to Stanovich and West (1998, 2000), information processing styles can be regarded as the second most fundamental predictor of individual differences in decision making. According to Evans and Over (2010), for instance, a theory of heuristic decision making even requires an underlying dual process framework of information processing to be complete. Dual process models of information processing have been widely applied (e.g., Cacioppo, Petty, Kao, & Rodriguez, 1986; Chaiken, 1980; Epstein, Pacini, Denes-Raj, & Heier, 1996; Evans & Over, 1996; Sloman, 1996; Tversky & Kahneman, 1983), thus showing the importance of this account. Although they have been formalized in slightly different ways, they all agree on the same basic idea of two fundamental information processing systems, one analytic-rational and one intuitive-experiential.[5] I focus on the Cognitive-Experiential Self-Theory (Epstein et al., 1996) as it offers a fruitful theoretical ground for deriving predictions regarding decision making with the RH. Furthermore and in contrast to virtually all other approaches, this theory also offers measurement tools for assessing individual differences with respect to the two proposed information processing styles. The corresponding findings are reported in section 4.3.

To assess potential determinants of RH-use, the application of powerful measures of RH-use is of special importance as it will boost the correlation with measures of personality (e.g., Stanovich & West, 2000). Thus, before I turn to the question of which person-specific factors might explain individual differences in heuristic use, it is important to briefly introduce some methodological aspects. These methodological preliminaries include the typical task environment as well as the newly developed measurement model of RH-use, which represents a main contribution to this thesis.

---

[5] Depending on the theory, the two systems of information processing are called differently, for instance, central and peripheral (Cacioppo et al., 1986), systematic and heuristic (Chaiken, 1980), rule-based and associative (Sloman, 1996), or extensional and intuitive (Tversky & Kahneman, 1983).

Individual differences in use of the RH

3. Methodological Background

"Clearly, the RH is a precise model which makes exact predictions about choices and underlying processes. However, to gain insight about whether and under which conditions these predictions are actually correct, measurement must also be precise."
(Hilbig, 2010, p. 272)

When investigating the RH, several conditions must be fulfilled (cf. Gigerenzer & Goldstein, 2011). First, the RH has been defined for natural recognition as opposed to experimentally induced recognition. However, this requirement has not always been fulfilled (e.g., Bröder & Eichler, 2006; Newell & Shanks, 2004; Oppenheimer, 2003). Second, inferences should be made from memory as opposed to inferences from givens, where information is openly presented to the decision maker. Also this proposition has not always been followed (e.g., Ayton et al., 2011; Bröder & Eichler, 2006; Goldstein & Gigerenzer, 2002; Newell & Fernandez, 2006; Richter & Späth, 2006). Third, the recognition validity should be substantial, rendering the RH a successful and thus useful strategy, as opposed to a value close to chance level (e.g., Newell & Shanks, 2004; Pohl, 2006; Richter & Späth, 2006). All empirical investigations reported in this thesis satisfied the concepts of natural recognition, memory-based knowledge, and RH-friendly decision domains as explained above.[6] There was only one exception from these preliminaries: A decision domain disfavoring RH-use was used in the second manuscript for specific reasons, as will be explained in section 4.2.

To investigate the RH, a paradigm has been established that consists of a recognition task and a paired comparison task (e.g., Goldstein & Gigerenzer, 2002; Pohl, 2006). To this end, a representative set of objects is randomly drawn from a reference class, that is, an exclusively defined set of objects (e.g., the world's 100 most populous cities with a population of more than 3 million inhabitants). In the recognition task, decision makers have to provide yes-no recognition judgments for each object. For the paired comparison task, the set of objects is (most often exhaustively) paired. In

---

[6] For sake of completeness, I would like to mention that further critical experimental set-ups have been discussed by Pachur et al. (2008), in particular, use of induced cue knowledge, use of criterion knowledge, presentation of cue knowledge about unrecognized objects, and use of artificial stimuli. However, none of these critical experimental set-ups is comprised in the studies reported here.

the comparison task, decision makers are asked to decide which of the two objects in each pair has the higher value with respect to the criterion of interest (e.g., which city is more populous). Again, all empirical investigations reported in this thesis adhere to this paradigm.

3.1. Assessing RH-Use

"How can one tell whether people are following the recognition heuristic or choosing correctly by some other means?" (Goldstein & Gigerenzer, 2002, p. 83)

Besides the methodological preliminaries and experimental set-ups described above, it is important to measure RH-use in an unbiased way, that is, without confounds with other strategies. Therefore, choice of the appropriate measure is discussed next.

3.1.1. Ad-Hoc Measures of RH-Use

In the last two decades, several measures of RH-use have been suggested (Hilbig & Pohl, 2008; Pachur & Hertwig, 2006). However, all of them display severe limitations, which are briefly sketched in the following. For an extensive review and comparison of all measures of RH-use described here, see Hilbig (2010) or Pachur (2011).

To begin with, the adherence rate is defined as the proportion of pairs where the recognized object is chosen. The central disadvantage of the adherence rate is that RH-use is confounded with knowledge integration. In other words, the recognized object can be chosen for different reasons, one being recognition alone, another being further knowledge retrieved from memory (e.g., Hilbig, 2010). Unfortunately, it is often not possible to disentangle different strategies based solely on choice patterns because these strategies may predict the same choices under certain circumstances. In particular, the adherence rate is not able to disentangle use of the RH and integration of knowledge when the available knowledge is in line with the recognition cue. Consequently, the adherence rate will typically overestimate the probability of RH-use (e.g., Hilbig & Richter, 2011).

Next, two measures derived from signal detection theory (SDT; for an introduction, see Macmillan & Creelman, 2005) have been proposed based on the SDT hit and false alarm rates (Pachur & Hertwig, 2006). When deciding between a recognized and an unrecognized object, choosing the recognized object denotes a hit, in case it represents the correct answer, and a false alarm, in case it represents the false answer. The index c displays the tendency to follow the recognition cue. It is formalized as $c = -\frac{1}{2}(z(H) + z(FA))$ with $z(H)$ and $z(FA)$ denoting the z-transformed hit and false alarm rates, respectively. The index d', formalized as $d' = z(H) - z(FA)$, displays the ability to discriminate cases where recognition leads to a correct versus a false inference. Certainly, the two indices provide less confounded measurements of RH-use than the adherence rate. But there are limitations to their interpretation as well. First, none of them can be understood as the proportion of RH-use. Second, it is not clear how to interpret the exact values of c and d'. In principle, true users of the RH should score $c < 0$ and $d' = 0$. However, under the realistic assumption of strategy execution errors, there are no conventions where to set the threshold for c, nor how wide to define the interval around zero for d' to classify decision makers as users versus non-users of the RH. Thus, c and d' can only be used as proxies for the probability of RH-use.

Finally, the discrimination index $DI = (H) - (FA)$ was introduced (Hilbig & Pohl, 2008), with (H) denoting the hit rate and (FA) the false alarm rate. In the same way as the index d', the DI is defined as the ability to discriminate cases where the recognition cue leads to a correct versus a false inference. There are two main differences between these two measures: First, the DI is not theoretically dependent on SDT. Second, it is defined based on hit and false alarm rates instead of their z-transformation. However, as the DI shares its interpretation with d', it also shares its advantages and disadvantages.

To overcome the limitations of the existing measures of RH-use, Hilbig, Erdfelder, and Pohl (2010) proposed the r-model, which is introduced in the next section. This model is of special interest for my thesis as it represents the basis of the newly developed measurement tool of RH-use.

3.1.2. The r-Model

The r-model (Hilbig, Erdfelder, & Pohl, 2010) belongs to the class of multinomial processing tree (MPT) models, that is, a class of statistical models that aim at explaining observed categorical data by latent cognitive processes (for an overview on MPT models, see Batchelder & Riefer, 1999; Erdfelder et al., 2009). As illustrated in Figure 2, the r-model is designed to account for the three cases that can occur in a paired comparison task: knowledge cases (both objects recognized), guessing cases (neither object recognized), and recognition cases (exactly one object recognized). In particular, responses to the paired comparison task are assigned to eight mutually exclusive categories. The category counts are assumed to follow a multinomial distribution and are accounted for by four latent parameters according to the model: recognition validity (parameter a), knowledge validity (parameter b), probability of correct guessing (parameter g), and probability of RH-use (parameter r). These model parameters can be estimated using maximum likelihood techniques (Hu & Batchelder, 1994), which allow for goodness-of-fit testing.

Consider the most relevant case: If one of the two objects is recognized, the RH is used with probability r, in which case the correct choice is made with the probability a, that recognition represents a valid cue. In contrast, additional knowledge is used with probability $1-r$. In this case, the correctness of the decision depends on the probability a, that recognition represents a valid cue, and additionally on the probability b, that the available knowledge is valid.

The r-model has been validated empirically and via simulations (Hilbig, 2010; Hilbig, Erdfelder, & Pohl, 2010). Here, I would like to draw attention to its two most important advantages. First, the model assesses RH-use free of confounds with knowledge integration (in contrast to the adherence rate). Second, the r parameter can be directly interpreted as the probability of RH-use and not only as a proxy (as opposed to the indices c, d', and DI). Thus, the r-model is superior to the measures of RH-use described before, but it displays limitations as well. These limitations are explained in more detail in the following section.
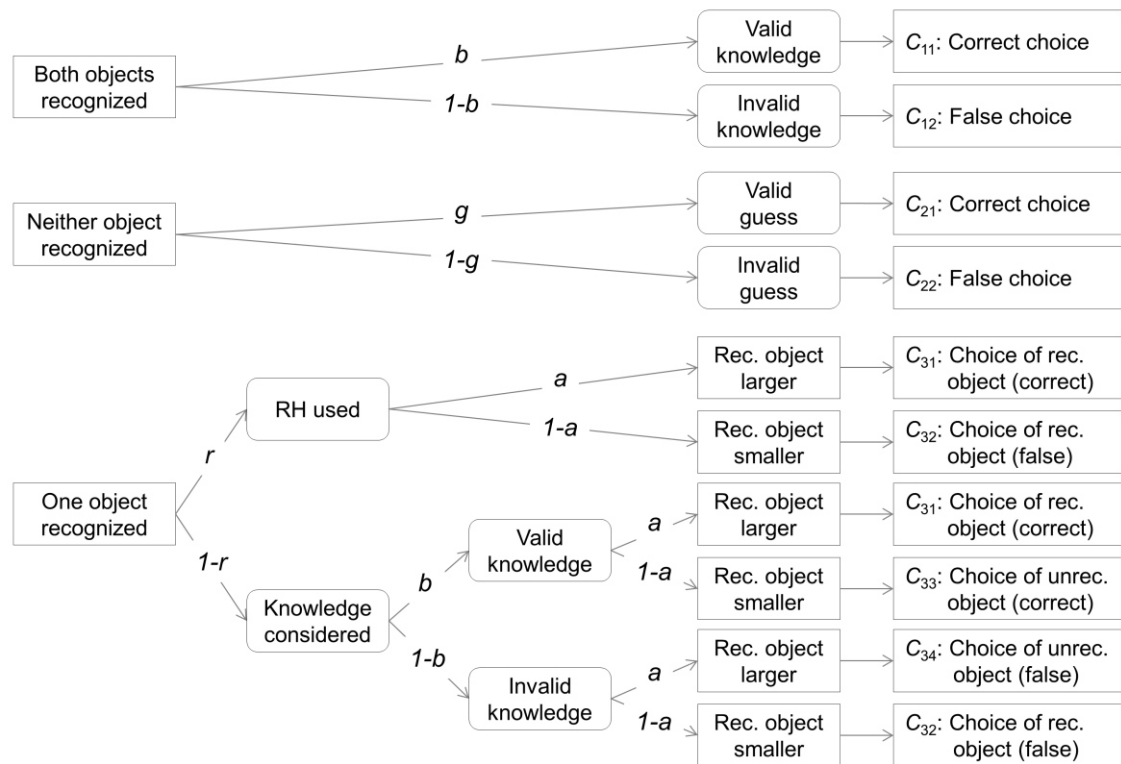
Figure 2. Illustration of the r-model (Hilbig, Erdfelder, & Pohl, 2010). The graph structure displays dependencies among observable events (rectangles) and the underlying latent states (rectangles with rounded corners). In particular, category counts for the eight mutually exclusive categories C_{ij} ($i \in \{1, 2, 3\}$, $j \in \{1, 2, 3, 4\}$) for the three object cases (i.e., both, neither, and one object recognized) are accounted for by the four model parameters a, b, g, and r, which represent recognition validity, knowledge validity, the probability of correct guessing, and the probability of RH-use, respectively.

3.2. Assessing Individual Differences in RH-Use

"Individual-level tests are essential because in virtually every task we find individual differences in strategies." (Gigerenzer & Brighton, 2009, p. 133)

Most studies on RH-use investigated strategy selection on the group level. However, group-level analyses are based on two problematic assumptions: independence of the underlying cognitive processes and homogeneity across individuals (e.g., Klauer, 2010; Matzke, Dolan, Batchelder, & Wagenmakers, 2015).

Stated differently, it is assumed that all aspects of task performance, that is, recognition validity, knowledge validity, guessing accuracy, and RH-use (as measured by the r-model parameters) are uncorrelated. It is further assumed that individuals do not differ regarding the processes estimated through the model parameters. The major problem is that violations of these two assumptions can lead to highly erroneous conclusions.

Unquestionably, relations between many cognitive processes do exist (e.g., LaBar, Gitelman, Parrish, & Mesulam, 1999; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010; Miyake et al., 2000). By implication, correlations between model parameters, representing the underlying cognitive processes, are very likely (Klauer, 2010; see also Matzke et al., 2015). Specifically, RH-use has been shown to increase with recognition validity, and to decrease with knowledge validity on the group level (see section 2.2). Consequently, these and other aspects of task performance might also be correlated on the individual level (e.g., Hilbig & Pohl, 2008; Pohl, 2006).

Similarly, individual variables can influence the performance in many cognitive tasks (e.g., Booth, Schinka, Brown, Mortimer, & Borenstein, 2006; Revelle, 1987; Rindermann & Neubauer, 2001). Accordingly, individual differences in personality traits, cognitive skills, or demographic factors may lead to violations of the homogeneity assumption of MPT models (e.g., Arnold, Bayen, & Böhm, 2014; Arnold, Bayen, & Smith, 2015; Coolin, Erdfelder, Bernstein, Thornton, & Thornton, 2015; Horn et al., 2015). As illustrated in section 2.3, this finding also holds for the RH and the r-model. In addition, a simple example demonstrates how grave this problem is: Individual recognition and knowledge validities might vary simply because individuals recognize different objects.

If the homogeneity assumption is violated, group-level analyses can result in biased parameter estimates, underestimated standard errors, and underestimated confidence intervals (e.g., Klauer, 2006; Riefer & Batchelder, 1991; Smith & Batchelder, 2008; Stahl & Klauer, 2007). As a consequence, goodness-of-fit tests and hypothesis tests will result in erroneous rejections of adequate models and overrated numbers of significant differences (e.g., Klauer, 2006; Smith & Batchelder, 2008). If the independence assumption is additionally violated, the extent of these statistical problems may vary widely depending on the covariance structure of the r-model parameters (Klauer, 2006, 2010). On the whole, given parameter heterogeneity and

Individual differences in use of the RH

correlations between parameters, group-level analyses can result in all kinds of deficient conclusions.

An alternative approach to assess RH-use, which accounts for heterogeneity across decision makers, is individual-level analysis, that is, applying the r-model to the data of each participant separately (e.g., Hilbig, Erdfelder, & Pohl, 2010, 2011; Hilbig & Richter, 2011). However, this procedure can also be problematic for several reasons. First, small numbers of observations per participant can lead to biased parameter estimates, large confidence intervals, and low power to detect model misfit (e.g., Chechile, 2009; Cohen, Sanborn, & Shiffrin, 2008; Hilbig, Erdfelder, & Pohl, 2010). To demonstrate, Hilbig, Erdfelder, and Pohl (2010) have shown that a minimum of 500 paired comparisons per individual is needed to reliably estimate the parameters of the r-model. However, studies on RH-use typically involve only about 90 to 190 comparisons per individual. As a consequence, individual-level analysis may lead to flawed results in most cases.

Second, the correlation between the hypothesized latent processes cannot be measured properly by the observed correlation of the estimated model parameters. To illustrate, the Pearson correlation often underestimates the true correlation between two latent variables when the observations are subject to measurement noise (e.g., Spearman, 1904). But measurement error represents the rule rather than an exception. In particular, the standard errors of the r parameter estimates, a proxy for measurement noise, typically range up to .25 on the individual level. Thus, correcting for the measurement error of the r-model's parameter estimates is of special importance when assessing correlations among parameters (Matzke et al., 2015). This will be the case in section 4.1.

Finally, the influence of external covariates on RH-use can only be assessed in a two-step procedure. In a first step, the r-model is applied to the data of individual persons. In a second step, the resulting estimates of RH-use are regressed on or correlated with personality test scores. As explained before, the regression or correlation analysis does not account for the measurement error of the r-model's parameter estimates and can therefore lead to biased estimates of the regression and correlation coefficients (e.g. Behseta, Berdyyeva, Olson, & Kass, 2009; Klauer, 2006; Matzke et al., 2015; Spearman, 1904). This problem especially accounts for the work presented in sections 4.2 and 4.3.

To overcome these problems, in recent years a growing number of researchers have proposed approaches to MPT models that accommodate heterogeneity between individuals. However, since this line of research is quite new, choice of the appropriate method has to be carefully considered. In the next section, I give a brief introduction to the most important approaches proposed so far and evaluate if they are appropriate to answer the research questions posed in this thesis.[7]

3.2.1. Hierarchical Approaches

In response to the issues described above, researchers have developed several hierarchical extensions to traditional MPT models. The basic idea of a hierarchical approach is that the model is specified on two levels. On the individual level, a separate and possibly different set of parameters is defined for each individual according to the basic MPT model (e.g., the r-model described in section 3.1.2). On the hierarchical group level, these individual parameters are assumed to follow some distribution, the so called hyper-distribution, which captures the variability between individuals using group-level parameters. Furthermore, to answer the research questions posed in this thesis, the hierarchical model should also account for the relations between parameters within the model as well as the relations to external trait variables.

To begin with, Klauer (2006) suggested hierarchical latent-class MPT models using a discrete hyper-distribution of group-level parameters. According to this approach, each participant can be assigned to one of the mutually exclusive latent classes of model parameters. Within each class, parameter homogeneity is assumed, whereas between classes there can be differences between parameters as well as correlations across parameters. Despite its ability to test and account for parameter heterogeneity, the latent-class approach features two major limitations. First, it can only

---

[7] In the following, I only consider heterogeneity between individuals as this represents the core interest of my thesis. Certainly, heterogeneity in stimulus materials might also exist. However, this problem might be less severe as stimulus materials are usually carefully selected and more thoroughly controlled than is possible with regard to participants. Indeed, there are no indications for an influence of particular stimulus objects on RH-use, neither in the studies this thesis is based on nor in the literature. Also, note that the methods described in the following sections can all be applied to situations where items (instead of individuals) differ. Additionally, a method to simultaneously handle heterogeneity in individuals and items has recently been developed by Matzke et al. (2015). Thus, it is principally also possible to investigate effects of stimulus materials on RH-use.

assess heterogeneity between an a-priori defined number of latent groups but not between individuals within these groups. To illustrate, although each class could theoretically consist of a single person, on a practical level the number of classes will be limited due to model identifiability concerns and to difficulties interpreting a large number of classes (Klauer, 2006). Second, the relation to external covariates, such as cognitive or personality traits, is not modeled. Therefore, the latent-class approach represents a better approximation than traditional MPT models only in specific situations (e.g., Stahl & Degner, 2007; Stahl & Klauer, 2008, 2009). For instance, it might be a useful tool when a sample of people can be split a-priori into homogeneous subclasses according to some key variable like education level or age groups. However, this is of little help if we are interested in the continuous effect of individual traits on RH-use. In this case, the application of continuous instead of discrete hyper-distributions seems more plausible.[8]

In contrast to Klauer's (2006) latent-class approach, Smith and Batchelder (2008) suggested the beta-MPT approach that captures the between-subject variability by means of continuous distributions, namely, beta distributions. However, as individual person parameters are assumed to follow independent beta distributions, the model does not account for correlations between parameters. Furthermore, until now, this approach does not incorporate the relations between external covariates and RH-use. Overall, the beta-MPT approach can indeed be useful for certain research questions (e.g., Arnold et al., 2014; Horn et al., 2015), but is not tailored to answer the questions posed in this thesis.

An approach incorporating all required features for my analyses is the hierarchical latent-trait approach to MPT models (Klauer, 2010). In particular, it accounts for heterogeneity within parameters and correlations across parameters. In

---

[8] An alternative non-hierarchical approach based on distinct classes or groups of individuals is recursive partitioning of MPT models (Strobl, Wickelmaier, & Zeileis, 2011; Wickelmaier & Zeileis, 2013). The main advantage compared to Klauer's (2006) latent-class approach is that the number of groups does not have to be determined a-priori but is defined based on the test scores of individual traits. However, just like the latent-class approach, it can only account for heterogeneity between groups, but not individuals within these groups.

addition, it allows for the assessment of the influence of cognitive and personality variables on RH-use in a straightforward way.[9]

3.2.2. The Latent-Trait r-Model

To investigate individual-level RH-use as well as its determinants, I am the first to propose a hierarchical extension of the r-model based on Klauer's (2010) latent-trait approach to MPT models. Thereby, I close a methodological gap in research on individual differences in RH-use. Compared to conventional analyses and to the hierarchical approaches described before, the latent-trait approach has several advantages (Klauer, 2010; Matzke et al., 2016). First, model parameters are estimated more reliably than by means of non-hierarchical analyses by using information from the group-level structure (i.e., the hyper-distribution). Thus, parameter estimates are more reliable even in case of small numbers of paired comparisons per individual. Second, by applying a multivariate distribution the latent-trait r-model not only accounts for correlations between model parameters but even explicitly models them. Third, the relation to external variables can be included into the model in terms of a regression. In this way, the latent-trait r-model allows for the estimation of model parameters, correlations between model parameters, and the influence of external variables on model parameters in a one-step procedure. Thereby, the estimated correlations and regressions are automatically adjusted for the uncertainty of the individual parameter estimates of the r-model. In other words, the latent-trait r-model is constructed to estimate the true underlying correlation between model parameters and the true regression of the model parameters on external variables, respectively, free of measurement error.

The latent-trait r-model, illustrated in Figure 3, is developed in two steps. In Michalkiewicz and Erdfelder (2016), we construct a hierarchical version of the r-model. Specifically, we extend it to account for two tests of the RH simultaneously. Thereby, we allow the assessment of the test-retest correlation, in other words, stability within

---

[9] Only recently, an alternative non-hierarchical approach to Klauer's (2010) latent-trait model has been suggested, which differs mainly by using a logit instead of a probit link function and by using Maximum Likelihood techniques instead of Bayesian statistics (Coolin et al., 2015). However, this account does not model correlations between parameters. Furthermore, model parameters are estimated less reliably within this non-hierarchical approach compared to the hierarchical model.

Individual differences in use of the RH

model parameters across two tests of the RH. In Michalkiewicz, Arden, and Erdfelder (2016) as well as Michalkiewicz, Minich, and Erdfelder (2016), we incorporate external covariates into the model to explain variation in RH-use through individual traits. Specifically, the model's parameters are estimated within the Bayesian framework. For this purpose, prior distributions are specified for all model parameters representing initial beliefs (as described in the caption of Figure 3). These prior distributions are then updated using the observed data. Thereby, posterior distributions are obtained whose statistical properties (in particular, the mean and the 95% Bayesian credible interval) are used to summarize the results. For a comprehensive introduction to hierarchical MPT models and more details on the Bayesian estimation process, see for instance Lee and Wagenmakers (2013).

The latent-trait r-model in Figure 3 shows the individual-level and the hierarchical group-level structure. In particular, the individual-level structure is based on the r-model. The group-level structure assumes that individual differences between participants are defined by a common continuous distribution – a multivariate normal distribution with mean vector $(\mu^a, \mu^b, \mu^g, \mu^r)$ and variance-covariance matrix Σ. The means μ^s ($s \in \{a, b, g, r\}$) are interpreted as the group level parameter estimates of recognition validity, knowledge validity, guessing validity, and probability of RH-use, respectively. Based on the variance-covariance matrix Σ, the standard deviations σ^s ($s \in \{a, b, g, r\}$) and the correlations $\rho_{s,t}$ ($s, t \in \{a, b, g, r\}$) between model parameters are derived. In particular, σ^s ($s \in \{a, b, g, r\}$) represent the variability across individuals, being close to zero when individuals are rather homogeneous and closer to one the larger the heterogeneity between individuals. For each individual i, $i = 1, …, I$, model parameters $s_i = \Phi(\mu^s + \xi^s \cdot \delta_i^s)$, $s \in \{a, b, g, r\}$, are modeled as the probit-transformed linear combination of the group mean μ^s, the individual displacement from the group mean δ_i^s, and a multiplicative scaling parameter ξ^s. To estimate the effect of an individual trait on RH-use, a regression term $\beta \cdot Cov_i$ is included in the linear function, consisting of a regression weight β and individual test scores of a trait variable Cov_i. Examples, on how to extend the model to account for the effect of two (or more) individual traits, are given in Michalkiewicz, Arden, and Erdfelder (2016), and Michalkiewicz, Minich, and Erdfelder (2016).
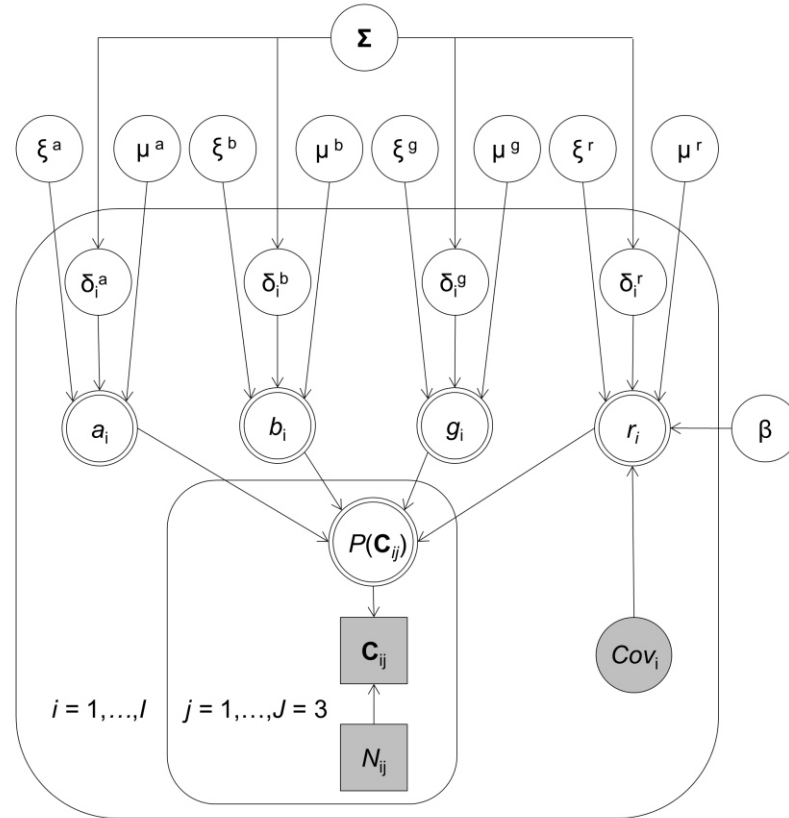
Individual differences in use of the RH

Figure 3. Illustration of the latent-trait r-model including a dispositional predictor of RH-use (adapted from Michalkiewicz & Erdfelder, 2016). The graph structure displays dependencies among observed data (shaded nodes) and latent model parameters (unshaded nodes). The square and circular nodes represent discrete and continuous variables, respectivley. The single-bordered nodes represent variables that have to be estimated from the data, while the double-bordered nodes represent variables that can be derived as a combination of other model parameters. The plates show the *I* individuals and the *J* object cases (i.e., both, neither, and one object recognized). Individual category probabilities $P(\mathbf{C}_{ij})$ are defined according to the r-model (Figure 2). Thereby, $\mathbf{C}_{ij} \sim$ Multinomial($P(\mathbf{C}_{ij})$, N_{ij}) represents the multinomially distributed category counts and N_{ij} the number of observations. Individual model parameters s_i ($s \in \{a, b, g, r\}$) are assessed as the probit-transformed linear combination of the group mean $\mu^s \sim$ Normal(0,1), a multiplicative scale parameter $\xi^s \sim$ Uniform(0,100), and individual displacement parameters $\delta_i^s \sim$ MultivariateNormal($\mathbf{0},\mathbf{\Sigma}$). The individual displacement parameters as well as the group standard deviations σ^s are derived from the covariance structure of the model's parameters $\mathbf{\Sigma} \sim$ Wishard($\mathbf{I}$,5). Additionally, the influence of an individual trait Cov_i on RH-use r_i is modeled in terms of a regression with regression coefficient $\beta \sim$ Normal(0,1).

4. Summaries of Manuscripts

"There are large individual differences in strategy selection. The attempt to find personality dimensions as correlates of strategy preferences has not been successful so far [...]." (Bröder & Newell, 2008, p. 208)

In the following sections, I provide summaries of the three manuscripts this thesis is based on. Thereby, I focus on the main results, achievements, and limitations of each manuscript. For the sake of brevity, I refrain from reiterating details as all information can be found in the original manuscripts located at the end of this thesis. An extensive discussion and outlook with regard to the overarching objective of this thesis is provided in section 5.

4.1. Stability in RH-Use

Michalkiewicz, M., & Erdfelder, E. (2016). Individual differences in use of the recognition heuristic are stable across time, choice objects, domains, and presentation formats. *Memory & Cognition, 44,* 454-468. doi:10.3758/s13421-015-0567-6

"It is yet an open question whether the different strategy preferences diagnosed in a one-shot assessment of an experiment will turn out to be stable across tasks and situations. If not then states rather than traits should be investigated as variables causing the individual differences [...]." (Bröder & Newell, 2008, p. 208)

Large individual differences in strategy selection have been consistently found in the literature, as I have summarized in section 2.3. However, before we can try to explain these differences in terms of cognitive or personality traits, an important condition needs to be checked in the first place: stability in the use of the RH. Only if RH-use represents a stable decision making style, it will be principally possible to find its personality determinants. For this purpose, we assessed stability in RH-use across time, choice objects, decision domains, and presentation formats. These situational

factors are assumed not to influence the overall level of RH-use as long as recognition and knowledge validities stay constant. By eliminating potential effects of situational factors we thus enabled the assessment of individual differences without confounds. Importantly, we estimated RH-use with all measures previously employed in the relevant literature (described in section 3) showing that stability in RH-use is not tied to statistical peculiarities of a single measurement tool (cf. Kantner & Lindsay, 2012). On the whole, we found temporal and cross-situational stability in RH-use similar in size compared to studies on other trait-like variables in judgment and decision making (e.g., Glöckner & Pachur, 2012; Kantner & Lindsay, 2012, 2014; Odum, 2011; Scheibehenne & Pachur, 2015; Witkin, Goodenough, & Karp, 1967). Specifically, as expected based on the review by Hilbig (2010), the adherence rate and the index c tended to overestimate stability in RH-use compared to the latent-trait r-model. In contrast, the DI and the index d' tended to underestimate it. This again shows the superiority of the latent-trait r-model compared to the remaining measures of RH-use.

Equally important, our results might also explain the difficulties to find dispositional determinants of RH-use. Aside from a small number of studies reporting such difficulties (e.g., Michalkiewicz, Coolin, & Erdfelder, 2013; Michalkiewicz, Hilbig, Erdfelder, Keller, & Bless, 2012; Pachur et al., 2009), it is highly probable that an even larger number of unpublished studies exists that show either inconclusive results or null-effects with regard to determinants of RH-use (cf. Appelt et al., 2011; Rosenthal, 1979). An attempt to explain these problems is based on the notion that the size of potential associations between RH-use and individual traits is limited by the size of stability in RH-use (cf. Mischel, 1968). In particular, the correlation between a powerful personality predictor and RH-use should not be expected to exceed the correlation between two tasks measuring RH-use, that is, stability in RH-use. Consequently, the correlation between a weak predictor and RH-use will be even lower and may be hardly detectable. Hence, the above described problems might be partly due to the investigation of weak predictors. So, the search for determinants of RH-use should be restricted to presumably powerful and theoretically well founded predictors, as already outlined in section 2.4.

At this point it should be noted that we could rule out a severe objection against our findings, namely that stability in RH-use is only an epiphenomenon of stability in individual recognition or knowledge validities. To reiterate, it has repeatedly been

shown that higher recognition validity leads to more RH-use, whereas more valid knowledge leads to less RH-use (see section 2.2). Therefore, one might hypothesize that stability in RH-use is caused by stable underlying differences in individual recognition and knowledge validities. To rule out this objection, we demonstrate 1) that stability in recognition and knowledge validities were both either comparable in size or even smaller than stability in RH-use; 2) that the correlation between RH-use and recognition or knowledge validity, respectively, was not reliable across individuals in nearly all studies; and 3) that stability in RH-use was not affected by partialling out the effect of recognition and knowledge validity, respectively. These results suggest that RH-use truly reflects a specific style of decision making. Now that stability is shown, I turn to the investigation of individual determinants of RH-use.

4.2. Intelligence and Adaptive RH-Use

Michalkiewicz, M., Arden, K., & Erdfelder, E. (2016). *Do smarter people make better decisions? The influence of intelligence on adaptive use of the recognition heuristic.* Manuscript submitted for publication.

"Successful decision making requires the selection of strategies that match the specific characteristics of the current decision task as well as the resources available to the decision maker." (Pachur et al., 2009, p. 902)

On the one hand, it has repeatedly been shown that decision makers successfully adapt their decision strategies to the situation, as outlined in section 2.2. On the other hand, it has consistently been found that individuals differ with respect to strategy selection, as shown in section 2.3. But one important question remained largely unanswered so far: Do individuals systematically differ in their ability to successfully adapt their decision strategies to different situations? And if so, which cognitive, demographic, or personality factors might explain these differences?

To answer these questions, we investigated intelligence as a potential determinant of adaptive RH-use. This idea is derived from the notion that intelligence represents the general cognitive capacity responsible for successful adaptation (e.g.,

Neisser et al., 1996; Sternberg & Salter, 1982). Additionally, we studied whether fluid and crystallized intelligence – the two major factors of general intelligence - offer an explanation for this effect. To illustrate, fluid intelligence has been conceived as the capability to understand complex relationships and to solve new problems (Horn & Cattell, 1966). Thus, fluid intelligence might affect the capability to analyze the environmental structure, that is, the size of the recognition and knowledge validities, and to select the most appropriate decision strategy accordingly, that is, RH-use versus knowledge integration. Crystallized intelligence reflects skills and expertise acquired through personal experience (Horn & Cattell, 1966). In this way, crystallized intelligence might also affect the ability to understand the fit between the environmental structure and the corresponding decision strategies. However, this effect will probably depend on whether or not an individual possesses prior task knowledge or experience with similar decision situations (Alba & Hutchinson, 1987; Chi, Glaser, & Farr, 1988; Mata, Schooler, & Rieskamp, 2007). In sum, both fluid and crystallized intelligence may affect successful adaptation to the decision context. Indeed, there is already initial evidence for effects of fluid and crystallized intelligence on adaptive heuristic use (Bröder, 2003; Mata et al., 2007; Pachur et al., 2009).

To this end, we first investigated a situation where the RH outperformed knowledge integration in terms of validity. Therefore, we reanalyzed the study by Hilbig (2008) that assessed general intelligence (among others) as a potential confounding variable when assessing the effect of neuroticism on RH-use. Second, we conducted an experiment comprising two conditions: one condition where knowledge integration outperformed RH-use in terms of validities as opposed to the situation investigated by Hilbig (2008), and a second condition where both strategies were equally well adapted to the decision context in terms of validities. Here, intelligence was assessed in terms of fluid and crystallized intelligence. On the whole, we expected intelligence to increase use of the smarter, meaning the more valid, strategy in case it is determined by the decision context. In contrast, we expected no influence of intelligence when the decision context is indifferent. Additionally, we analyzed whether this effect is mainly driven by fluid or crystallized intelligence.

In line with our hypotheses, in Hilbig's (2008) experiment general intelligence was positively associated with RH-use when the RH outperformed knowledge integration. Conversely, in our experiment, RH-use was negatively associated with

fluid intelligence when knowledge integration outperformed RH-use. In contrast, we found no reliable association between fluid intelligence and RH-use when both strategies were equally valid. However, crystallized intelligence could not be shown to influence adaptive RH-use in our study. In sum, our results suggest that intelligence, specifically fluid intelligence, moderates adaptivity in RH-use.

Notably, there are limitations to our study that could be responsible for the null-effect of crystallized intelligence. First, we used a test of crystallized intelligence that is designed for the general population and not specifically tailored to student samples. This might explain the somewhat limited range of crystallized intelligence in our student sample that perhaps contributed to the smaller effects for this intelligence measure. Second and related to this issue, the selected sample might have been too homogeneous to find a correlation, that is, the variance in intelligence was not sufficient in size. Third, one might argue that the tests that we applied to assess crystallized intelligence did not capture those aspects of crystallized intelligence which affect adaptive strategy selection (cf. Beauducel, Liepmann, Felfe, & Nettelnstroth, 2007). Fourth, it is possible that the inference tasks that we used were too artificial to elicit use of knowledge and skills acquired through everyday experience. All these aspects might have contributed to the small effect of crystallized intelligence.

Note that while the decision context did not favor any strategy, that is, neither knowledge nor the RH, we found a weak positive association between knowledge-use and fluid intelligence, at least descriptively. In other words, we found that more intelligent decision makers preferred knowledge-use over RH-use in this situation although there was no situation-related reason for such a preference. There is one factor which might offer a reasonable explanation for this finding, namely, need for cognition (NFC; Cacioppo & Petty, 1982). NFC is defined as the extent to which people engage in and enjoy cognitively demanding activities. In fact, one might argue that intelligent people prefer the more complex and cognitively demanding strategy of knowledge integration over simple RH-use because they like to engage in cognitively demanding activities and enjoy use of cognitively demanding decision strategies. Corroborating this idea, a positive correlation between (fluid) intelligence and NFC has repeatedly been reported in the literature (e.g., Cacioppo, Petty, Feinstein, & Jarvis, 1996; Fleischhauer et al., 2010; Furnham & Thorne, 2013; Hill et al., 2013; von Stumm, 2013). These ideas are further elaborated in the following section.

4.3. Need for Cognition, Faith in Intuition, and RH-Use

Michalkiewicz, M., Minich, B., & Erdfelder, E. (2016). *Explaining individual differences in fast-and-frugal decision making: The impact of need for cognition and faith in intuition on use of the recognition heuristic.* Manuscript submitted for publication.

"Individuals seem to have personal tendencies that favor the use of compensatory or non-compensatory decision strategies, which are based on personality traits [...]."
(Shiloh, Koren, & Zakay, 2001, p. 701)

In the following, I address the Cognitive-Experiential Self-Theory (Epstein et al., 1996). This theory represents a plausible theoretical framework for deriving predictions on decision makers' propensities to rely on fast-and-frugal versus cognitively demanding strategies (cf. Appelt et al., 2011; Mohammed & Schwall, 2009), in this case, RH-use versus knowledge integration. According to the Cognitive-Experiential Self-Theory, individuals can make use of two systems of information processing, which are independent but can work simultaneously. The analytic-rational system is assumed to operate through reasoning, formal logic, and abstract thought. Thus, it is expected to be slow and effortful. The intuitive-experiential system, by contrast, is assumed to operate through personal experience, categorical thinking, and concrete examples. So, it is expected to be fast and effortless. The theory predicts that in practice everybody uses both processing systems depending on the situation and the decision task at hand, but has a preference for one of them (e.g., Phillips et al., 2016).

Preferences for or against these two systems of information processing can be assessed by means of two independent personality traits (Epstein et al., 1996): need for cognition (NFC) and faith in intuition (FII). NFC is defined as the extent to which a person likes to engage in effortful cognitive tasks, while FII is defined as the extent to which a person trusts in her or his intuitive feelings and immediate impressions. Indeed, there is already evidence that high NFC is associated with less use of certain heuristics, like the anchoring and adjustment heuristic (Carnevale, Inbar, & Lerner, 2011; Epley & Gilovich, 2006). In contrast, high FII has been shown to be associated with enhanced use of certain heuristics, like the ease-of-retrieval, the representativeness, and the

reinforcement heuristic (Alós-Ferrer & Hügelschäfer, 2012; Danziger, Moran, & Rafaely, 2006; Mahoney, Buboltz, Levin, Doverspike, & Svyantek, 2011). However, there is also evidence that these effects might not hold for all kinds of heuristic decision making (Bröder, 2012; Keller, Bohner, & Erb, 2000).

Based on these insights, we hypothesized that NFC counteracts RH-use while FII fosters it. In particular, we investigated the effects of NFC and FII in two conditions in line with the idea of strong and weak decision situations (Mischel, 1973; see also Lozano, 2016). In the strong condition the RH outperformed knowledge integration in terms of validity, whereas in the weak condition both strategies were equally well adapted to the decision context. We investigated the strong condition because it resembles the standard set-up used in most studies on the RH. Here, the bias towards RH-use might interfere with potential effects of personality and thereby hamper detection of such effects. In contrast, the weak condition was not biased towards any of the decision strategies. By implication, it should facilitate detection of potential effects of NFC and FII on strategy selection. By this means, we were able to interpret whether potential effects are only small (i.e., visible only in the weak condition) or can be considered substantial (i.e., persisting also in the strong condition).

As expected, we found a negative effect of NFC in both conditions. Notably, the effect was very similar in size across conditions. In other words, the negative effect of NFC became apparent not only in the weak condition, where situational influences are minimal, but also in the strong condition, where situational influences are high. This suggests that NFC can be considered a powerful predictor of RH-use in the sense that it was not reduced by situational influences (Lozano, 2016). In particular, the effect of NFC can be interpreted as high compared to the test-retest correlations found by Michalkiewicz and Erdfelder (2016).

Contrary to our expectations, our results do not show evidence for a positive effect of FII. This lack of evidence does not seem to be particularly uncommon, suggesting that the association between FII and the preference for heuristic use is less clear than predicted by Epstein et al. (1996) (e.g., Betsch & Glöckner, 2010; Keller et al., 2000; Pachur & Spaar, 2015). Apart from that, there are also alternative explanations for this lack of evidence. First, the effect of FII might have been too small to be detected within a relatively homogenous sample of students. Corroborating this idea, the range of FII scores was quite limited in our study. This might have resulted in

an underestimation of the true underlying effect of FII. Second, the effect of NFC was potentially masking the effect of FII. In line with this explanation, NFC was rather high in our study, whereas the FII value was rather low compared to previous studies (e.g., Epstein et al., 1996; Pacini & Epstein, 1999). Finally, it has been shown that decision makers tend to apply an intuitive strategy when they have substantial experience with a decision situation (e.g., Betsch, 2008; Pachur & Marinello, 2013), which might not have been the case in our experiment.

At this point, I would like to discuss our results in light of the experiment by Hilbig, Scholl, and Pohl (2010). When investigating the situational effect of processing modes, Hilbig, Scholl, and Pohl (2010) found that participants who were instructed to think deliberately applied the RH more often compared to those who were instructed to think intuitively. By contrast, we assessed the effect of processing styles as an individual predisposition and found the opposite pattern, namely that those preferring a deliberate processing style used the RH less often. The following line of thoughts might reconcile these opposed findings: Individuals who were instructed to use deliberate processing but principally favor intuitive processing may have been unable to cope with the cognitive effort inherent in knowledge integration (Cacioppo & Petty, 1982; Phillips et al., 2016). Therefore, they probably tried to compensate this cognitive effort by using a less demanding strategy, that is, the RH. In contrast, individuals who were instructed to use intuitive processing but principally prefer deliberate processing may have avoided applying the RH. Instead, they probably used their knowledge because they enjoy working on cognitively demanding tasks and also enjoy the cognitive effort associated with information integration (Cacioppo & Petty, 1982). One could assume that the result observed by Hilbig, Scholl, and Pohl (2010) may be due to a subsample of participants low in NFC. Unfortunately, since measures of NFC were not assessed in their study, there is no way to directly test these speculations.

On the whole, we found supportive evidence that NFC is negatively related to RH-use, in line with the idea that high NFC is associated with enhanced elaboration (Dole & Sinatra, 1998). However, we did not find converging evidence that FII is positively related to RH-use, contrary to our expectations that high FII is associated with affective and associative information processing (Zimmerman, Redker, & Gibson, 2011). To conclude, NFC partly explains individual differences in RH-use irrespective of any influence of the decision context.

5. General Discussion and Outlook

"It is an important quest to uncover the conditions and individual differences which foster or hamper application of simple one-cue strategies, such as the RH." (Hilbig, 2010, p. 282)

The overarching aim of my thesis is to explain the large variation in decision strategy selection with respect to the adaptive toolbox, in particular to the RH. To this end, my thesis fulfills three major goals. The first is a methodological goal, namely, to establish a tool for assessing individual RH-use and its determinants. In fact, the quest for precise tests of the RH started a decade ago by Pachur and Hertwig (2006) and Hilbig and Pohl (2008), and was further pursued by Hilbig, Erdfelder, and Pohl (2010). These researchers all introduced more elaborate measures of RH-use than the simple adherence rate. However, until now, there was no method to assess the dispositional factors underlying individual differences in RH-use in an unbiased and straightforward way. Thus, in response to the call for precise measures, I extended the r-model (Hilbig, Erdfelder, & Pohl, 2010). This model can not only be regarded as the standard method for investigating RH-use, but also serves as a fruitful foundation for other approaches analyzing different aspects of the RH (e.g., Castela et al., 2014; Heck & Erdfelder, in press). In particular, I created a hierarchical measurement model based on Klauer's (2010) latent-trait approach to MPT models, representing state of the art methodology. Specifically, this model allows less error-prone estimation of RH-use than previous methods. Moreover, it enables identification of the individual traits that explain variation in RH-use. Thus, researchers are now equipped with a helpful tool for assessing individual-level RH-use and its determinants.

My second goal, which is more of a theoretical nature, is the demonstration of temporal and cross-situational stability in RH-use. In fact, stability represents an essential precondition in any investigation of individual differences (Aminoff et al., 2012; Couch & Keniston, 1960; Kantner & Lindsay, 2012, 2014; Odum, 2011; Scheibehenne & Pachur, 2015; Yechiam & Busemeyer, 2008; Witkin et al., 1967). Specifically, within the adaptive toolbox, the need for stability assessment has been pointed out by Bröder and Newell (2008). Only if stability in RH-use can be shown, it will be possible to find substantial and replicable relations to individual trait variables.

In sum, by demonstrating temporal and cross-situational stability, my work provides the theoretical basis for the research on the dispositional determinants of RH-use.

The third is an empirical goal, which builds on the former two. Equipped with the appropriate method and after demonstrating that there is stability in RH-use, it is possible to investigate the effects of individual traits on RH-use (and also the interaction of individual traits with the decision context). The importance of this quest has been emphasized by both the decision making field (e.g., Payne et al., 1993) and personality psychology (e.g. Funder, 2001). In brief, I demonstrate that RH-use is not only influenced by properties of the decision task and the decision context, but also by individual cognitive and personality traits. So, my work contributes to a new line of research that focuses on individual differences in RH-use and that aims at explaining these differences in terms of personality traits as well as in terms of personality-situation interactions.

A point of debate concerning my thesis might be the application of a model resting upon Bayesian parameter estimation. In particular, a common critique to Bayesian statistics is that it is based on prior beliefs. The prior distributions, representing these initial beliefs, are often supposed to be arbitrarily chosen and to strongly influence the results. To rule out the first objection, we applied priors that have been established by Matzke, Lee, and Wagenmakers (2013) and that have been repeatedly used (e.g., Arnold et al., 2015; Matzke et al., 2015). Also, we openly report our choice to enable all readers to draw their own conclusions independently. Contradicting the second objection, note that Bayesian parameter estimation has been shown not to be sensitive to the choice of the prior distributions as long as sufficiently informative data are available (e.g., Agresti, Caffo, & Ohman-Strickland, 2004; Matzke et al., 2015). Stated differently, enough data can minimize or even eliminate the influence of the chosen priors. Given the fairly large sample sizes (ranging between 70 and 135 participants per study) and the large number of comparisons per individual (being 300 for all reported studies) we certainly met the criterion of relying on sufficiently informative data. Corroborating this assumption, a replication of the analyses with the basic r-model and standard correlation and regression methods showed the same patterns of results, only overall smaller in size. To conclude, both common objections to Bayesian statistics can be shown to be invalid here.

Individual differences in use of the RH

Where to go on from here? First of all, the reported studies should be repeated with more heterogeneous populations of participants than students. This endeavor is of special importance as we cannot show reliable effects of crystallized intelligence and FII on strategy selection. This lack of evidence might be due to the limited range of test scores in the homogeneous student samples. Arguing along the same lines, it is also possible that the effects of fluid intelligence and NFC are underestimated using student samples. Consequently, our effect sizes potentially only show lower bounds of the true effects with regard to the whole population. Use of more heterogeneous, thus more representative samples, would render the results more generalizable.

As a second step, one could think of other determinants of RH-use, alone or in combination, as well as further interaction effects of individual traits with the decision context. However, one should not expect all cognitive and personality traits to be relevant predictors of strategy selection. It is therefore important to select individual traits that are theoretically well-grounded (e.g., Appelt et al., 2011; Mohammed & Schwall, 2009). First, it would be interesting to study the effect of fluid intelligence and NFC in combination because relations between these two factors have repeatedly been demonstrated (as summarized in section 4.2). This way, we could find out if one of the effects is influenced or even explained by the other. Apart from this, impulsivity seems to be a reasonable predictor of RH-use, as impulsivity and RH-use are both associated with fast decision making (e.g., Brunas-Wagstaff, Bergquist, & Wagstaff, 1994; Dickman & Meyer, 1988; Mann, 1973; Pachur & Hertwig, 2006; Pachur et al., 2009; Volz et al., 2006). Even if impulsivity by itself has not been shown to predict use of the TTB heuristic (Bröder, 2012), and might consequently also show no influence use of the RH, a context condition such as time pressure may turn it into one. This would be in line with the idea that certain personality traits possibly reveal their influence only under certain context conditions (e.g., Appelt et al., 2011; Mohammed & Schwall, 2009) – similar to the effect of intelligence shown in section 4.2. Extending the idea of traits and corresponding situational contexts, one could examine the effect of neuroticism in more depth. In particular, Hilbig (2008) suggested that subjects high in neuroticism prefer RH-use over knowledge-use to avoid a diagnostic test of their abilities. If this assumption holds, increasing the personal relevance of the task may boost the effect of neuroticism. These ideas represent only a small extract and certainly more could be listed.

An important third step would be transferring the results to other fast-and-frugal heuristics. For instance, the TTB heuristic and the RH share recognition information as the basis for decisions. In fact, the RH even represents a building block of TTB. As a consequence, the results found here should also account for the TTB heuristic. However, note that the investigation of personality determinants of TTB has shown to be difficult (e.g., Bröder, 2012). One explanation might be that there are no measurement models of the TTB heuristic that estimate the probability of TTB-use. Until now, only classification methods exist that group individuals into users of TTB versus users of other strategies on the basis of their decision patterns. Thus, more elaborate methods are needed before starting this line of research. Another explanation might be that stability in use of TTB, an important precondition, has not been demonstrated yet. It seems likely that use of the TTB heuristic will be stable to a similar degree as use of the RH. However, as the TTB heuristic is more complex than the RH, without empirical tests we cannot be sure that individuals will indeed show stable propensities to rely on this strategy. Beyond that, transferring the results to even more extensive decision strategies involving the RH principle will probably also reveal problems. Certainly, this will require additional theoretical and empirical work.

Finally, it would be interesting to test whether the preference for RH-use is associated with certain response tendencies and biases typically found in judgment and decision making. A temporal and cross-situational stable preference for certain decision strategies in combination with a liability to certain response tendencies and biases could then be interpreted as evidence for a general heuristic decision-making style (Kantner & Lindsay, 2012). Picking up this idea, it would be interesting to explore if there is any individual trait variable, which manifests not only in a preference for the RH but also for other heuristics, response tendencies and biases.

To conclude, my thesis represents not only an important step in developing methods for analyzing individual differences in RH-use but also an important step in uncovering their sources. Thus, my thesis broadens our understanding of the RH by adding to a comprehensive framework that incorporates situational and individual determinants of strategy use, as well as their interaction. Thereby, it hopefully inspires future work on individual differences in strategy selection, their contextual and dispositional determinants, and the methods used to accomplish these tasks within the theory of the adaptive toolbox and beyond.

6. References

Agresti, A., Caffo, B., & Ohman-Strickland, P. (2004). Examples in which misspecification of a random effects distribution reduces efficiency, and possible remedies. *Computational Statistics and Data Analysis, 47*, 639-653. doi:10.1016/j.csda.2003.12.009

Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of consumer expertise. *The Journal of Consumer Research, 13,* 411-454. doi:10.1086/209080

Alós-Ferrer, C., & Hügelschäfer, S. (2012). Faith in intuition and behavioral biases. *Journal of Economic Behavior & Organization*, *84*, 182-192. doi:10.1016/j.jebo.2012.08.004

Aminoff, E. M., Clewett, D., Freeman, S., Frithsen, A., Tipper, C., Johnson, A., . . . Miller, M. B. (2012). Individual differences in shifting decision criterion: A recognition memory study. *Memory & Cognition*, *40*, 1016-1030. doi:10.3758/s13421-012-0204-6

Appelt, K. C., Milch, K. F., Handgraaf, M. J. J., & Weber, E. U. (2011). The decision making individual differences inventory and guidelines for the study of individual differences in judgment and decision-making research. *Judgment and Decision Making, 6,* 252-262. Retrieved from http://journal.sjdm.org/11/11218/jdm11218.pdf

Arnold, N. R., Bayen, U. J., & Böhm, F. B. (2014). Is prospective memory related to depression and anxiety? A hierarchical MPT modelling approach. *Memory, 23,* 1215-1228. doi:10.1080/09658211.2014.969276

Arnold, N. R., Bayen, U. J., & Smith, R. E. (2015). Hierarchical multinomial modeling approaches: An application to prospective memory and working memory. *Experimental Psychology, 62,* 143-152. doi:10.1027/1618-3169/a000287

Ayton, P., Önkal, D., & McReynolds, L. (2011). Effects of ignorance and information on judgments and decisions. *Judgment and Decision Making*, *6*, 381-391. Retrieved from http://journal.sjdm.org/11/rh7/rh7.html

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57-86. doi:10.3758/BF03210812

Beauducel, A., Liepmann, D., Felfe, J., & Nettelnstroth, W. (2007). The impact of different measurement models for fluid and crystallized intelligence on the correlation with personality traits. *European Journal of Psychological Assessment, 23,* 71-78. doi:10.1027/1015-5759.23.2.71

Behseta, S., Berdyyeva, T., Olson, C. R., & Kass, R. E. (2009). Bayesian correction for attenuation of correlation in multi-trial spike count data. *Journal of Neurophysiology, 101*, 2186-2193. doi:10.1152/jn.90727.2008

Betsch, C. (2008). Chronic preferences for intuition and deliberation in decision making: Lessons learned about intuition from an individual differences approach. In H. Plessner, C. Betsch, & T. Betsch (Eds.), *Intuition in judgment and decision making* (pp. 231-248). Mahwah, NJ: Lawrence Erlbaum.

Betsch, T., & Glöckner, A. (2010). Intuition in judgment and decision making: Extensive thinking without effort. *Psychological Inquiry, 21*, 279-294. doi:10.1080/1047840X.2010.517737

Booth, J. E., Schinka, J. A., Brown, L. M., Mortimer, J. A., & Borenstein, A. R. (2006). Five-factor personality dimensions, mood states, and cognitive performance in older adults. *Journal of Clinical and Experimental Neuropsychology*, *28,* 676-683. doi:10.1080/13803390590954209

Borges, B., Goldstein, D. G., Ortmann, A., & Gigerenzer, G. (1999). Can ignorance beat the stock market? Name recognition as a heuristic for investing. In G. Gigerenzer, P. M. Todd, and the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 59-72). New York, NY: Oxford University Press.

Bröder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 611-625. doi:10.1037/0278-7393.29.4.611

Bröder, A. (2012). The quest for take the best – Insights and outlooks from experimental research. In P. M. Todd, G. Gigerenzer, & the ABC Research Group (Eds.), *Ecological rationality: Intelligence in the world* (pp. 216-240). New York, NY: Oxford University Press.

Bröder, A., & Eichler, A. (2006). The use of recognition information and additional cues in inferences from memory. *Acta Psychologica, 121*, 275-284. doi:10.1016/j.actpsy.2005.07.001

Bröder, A., & Newell, B. R. (2008). Challenging some common beliefs: Empirical work within the adaptive toolbox metaphor. *Judgment and Decision Making*, *3*, 205-214. Retrieved from http://journal.sjdm.org/bn2.pdf

Brunas-Wagstaff, J., Bergquist, A., & Wagstaff, G. F. (1994). Cognitive correlates of functional and dysfunctional impulsivity. *Personality and Individual Differences*, *17*, 289-292. doi:10.1016/0191-8869(94)90033-7

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 116-131. doi:10.1037/0022-3514.42.1.116

Cacioppo, J. T., Petty, R. E., Feinstein, J., & Jarvis, W. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*, 197-253. doi:10.1037/0033-2909.119.2.197

Cacioppo, J. T., Petty, R. E., Kao, C. F., & Rodriguez, R. (1986). Central and peripheral routes to persuasion: An individual difference perspective. *Journal of Personality and Social Psychology*, *51*, 1032-1043. doi:10.1037/0022-3514.51.5.1032

Carnevale, J. J., Inbar, Y., & Lerner, J. S. (2011). Individual differences in need for cognition and decision-making competence among leaders. *Personality and Individual Differences*, *51*, 274-278. doi:10.1016/j.paid.2010.07.002

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytical studies.* New York, NY: Cambridge University Press.

Castela, M., Kellen, D., Erdfelder, E., & Hilbig, B. E. (2014). The impact of subjective recognition experiences on recognition heuristic use: A multinomial processing tree approach. *Psychonomic Bulletin & Review, 21*, 1131-1138. doi:10.3758/s13423-014-0587-4

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*, 752-766. doi:10.1037/0022-3514.39.5.752

Chechile, R. A. (2009). Pooling data versus averaging model fits for some prototypical multinomial processing tree models. *Journal of Mathematical Psychology, 53*, 562-576. doi:10.1016/j.jmp.2009.06.005

Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, A. L., Sanborn, A. N., & Shiffrin, R. M. (2008). Model evaluation using grouped or individual data. *Psychonomic Bulletin & Review, 15*, 692-712. doi:10.3758/PBR.15.4.692

Coolin, A., Erdfelder, E., Bernstein, D. M., Thornton, A. E., & Thornton, W. L. (2015). Explaining individual differences in cognitive processes underlying hindsight bias. *Psychonomic Bulletin & Review, 22*, 328-348. doi:10.3758/s13423-014-0691-5

Couch, A., & Keniston, K. (1960). Yeasayers and naysayers: Agreeing response set as a personality variable. *The Journal of Abnormal and Social Psychology*, *60*, 151-174. doi:10.1037/h0040372

Danziger, S., Moran, S., & Rafaely, V. (2006). The influence of ease of retrieval on judgment as a function of attention to subjective experience. *Journal of Consumer Psychology*, *16*, 191-195. doi:10.1207/s15327663jcp1602_9

Dickman, S. J., & Meyer, D. E. (1988). Impulsivity and speed-accuracy tradeoffs in information processing. *Journal of Personality and Social Psychology*, *54*, 274-290. doi:10.1037/0022-3514.54.2.274

Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist*, *33*, 109-128. doi:10.1080/00461520.1998.9653294

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science*, *17*, 311-318. doi:10.1111/j.1467-9280.2006.01704.x

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology*, *71*, 390-405. doi:10.1037/0022-3514.71.2.390

Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models. *Zeitschrift für Psychologie - Journal of Psychology, 217,* 108-124. doi:10.1027/0044-3409.217.3.108

Erdfelder, E., Küpper-Tetzel, C. E., & Mattern, S. D. (2011). Threshold models of recognition and the recognition heuristic. *Judgment and Decision Making, 6*, 7-22. Retrieved from http://journal.sjdm.org/11/rh13/rh13.pdf

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning.* Hove, UK: Psychology Press.

Evans, J. St. B. T., & Over, D. E. (2010). Heuristic thinking and human intelligence: A commentary on Marewski, Gaissmaier and Gigerenzer. *Cognitive Processing, 11,* 171-175. doi:10.1007/s10339-009-0339-y

Fleeson, W., & Noftle, E. (2008). The end of the person-situation debate: An emerging synthesis in the answer to the consistency question. *Social and Personality Psychology Compass, 2,* 1667-1684. doi:10.1111/j.1751-9004.2008.00122.x

Fleischhauer, M., Enge, S., Brocke, B., Ullrich, J., Strobel, A., & Strobel, A. (2010). Same or different? Clarifying the relationship of need for cognition to personality and intelligence. *Personality and Social Psychology Bulletin*, *36*, 82-96. doi:10.1177/0146167209351886

Frosch, C. A., Beaman, C. P., & McCloy, R. (2007). A little learning is a dangerous thing: An experimental demonstration of ignorance-driven inference. *The Quarterly Journal of Experimental Psychology*, *60*, 1329-1336. doi:10.1080/17470210701507949

Funder, D. C. (2001). Personality. *Annual Review of Psychology, 52,* 197-221. doi:10.1146/annurev.psych.52.1.197

Funder, D. C. (2008). Persons, situations, and person-situation interactions. In O. P. John, R. W. Richard, and L. A. Pervin (Eds.), *Handbook of Personality: Theory and Research* (3rd ed., pp. 568-580). New York, NY: Guilford Press.

Furnham, A., & Thorne, J. D. (2013). Need for cognition: Its dimensionality and personality and intelligence correlates. *Journal of Individual Differences*, *34*, 230-240. doi:10.1027/1614-0001/a000119

Gaissmaier, W., & Marewski, J. N. (2011). Forecasting elections with mere recognition from small, lousy samples: A comparison of collective recognition, wisdom of crowds, and representative polls. *Judgment and Decision Making, 6,* 73-88. Retrieved from http://nbn-resolving.de/urn:nbn:de:bsz:352-279320

Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. In D. J. Koehler and N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making* (pp. 62-88). Malden, MA: Blackwell Publishing Ltd.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*, 107-143. doi:10.1111/j.1756-8765.2008.01006.x

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451-482. doi:10.1146/annurev-psych-120709-145346

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The take the best heuristic. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 75-95). New York, NY: Oxford University Press.

Gigerenzer, G., & Goldstein, D. G. (2011). The recognition heuristic: A decade of research. *Judgment and Decision Making, 6,* 100-121. Retrieved from http://journal.sjdm.org/11/rh15/rh15.pdf

Gigerenzer, G., & Selten, R. (2001). *Bounded rationality: The adaptive toolbox.* Cambridge, MA: MIT Press.

Gigerenzer, G., Todd, P. M., & the ABC Research Group (1999). *Simple heuristics that make us smart.* New York, NY: Oxford University Press.

Glöckner, A., & Bröder, A. (2011). Processing of recognition information and additional cues: A model-based analysis of choice, confidence, and response time. *Judgment and Decision Making*, *6*, 23-42. Retrieved from http://journal.sjdm.org/11/rh4/rh4.pdf

Glöckner, A., & Bröder, A. (2014). Cognitive integration of recognition information and additional cues in memory-based decisions. *Judgment and Decision Making*, *9*, 35-50. Retrieved from http://journal.sjdm.org/13/13912/jdm13912.html

Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition*, *123*, 21-32. doi:10.1016/j.cognition.2011.12.002

Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 37-58). New York, NY: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109,* 75-90. doi:10.1037/0033-295X.109.1.75

Goldstein, D. G., & Gigerenzer, G. (2009). Fast and frugal forecasting. *International Journal of Forecasting, 25,* 760-772. doi:10.1016/j.ijforecast.2009.05.010

Heck, D. W., & Erdfelder, E. (in press). Extending multinomial processing tree models to measure the relative speed of cognitive processes. *Psychonomic Bulletin & Review*.

Hertwig, R., & Pachur, T. (2015). Heuristics, history of. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., pp. 829-835). Oxford, UK: Elsevier.

Hilbig, B. E. (2008). Individual differences in fast-and-frugal decision making: Neuroticism and the recognition heuristic. *Journal of Research in Personality, 42,* 1641-1645. doi:10.1016/j.jrp.2008.07.001

Hilbig, B. E. (2010). Precise models deserve precise measures: A methodological dissection. *Judgment and Decision Making, 5,* 272-284. Retrieved from http://journal.sjdm.org/10/rh5/rh5.pdf

Hilbig, B. E., Erdfelder, E., & Pohl, R. F. (2010). One-reason decision making unveiled: A measurement model of the recognition heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 123-134. doi:10.1037/a0017518

Hilbig, B. E., Erdfelder, E., & Pohl, R. F. (2011). Fluent, fast, and frugal? A formal model evaluation of the interplay between memory, fluency, and comparative judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 827-839. doi:10.1037/a0022638

Hilbig, B. E., Erdfelder, E., & Pohl, R. F. (2012). A matter of time: Antecedents of one-reason decision making based on recognition. *Acta Psychologica, 141*, 9-16. doi:10.1016/j.actpsy.2012.05.006

Hilbig, B. E., Michalkiewicz, M., Castela, M., Pohl, R. F., & Erdfelder, E. (2015). Whatever the cost? Information integration in memory-based inferences depends on cognitive effort. *Memory & Cognition, 43*, 659-671. doi:10.3758/s13421-014-0493-z

Hilbig, B. E., & Pohl, R. F. (2008). Recognizing users of the recognition heuristic. *Experimental Psychology, 55*, 394-401. doi:10.1027/1618-3169.55.6.394

Hilbig, B. E., & Pohl, R. F. (2009). Ignorance- versus evidence-based decision making: A decision time analysis of the recognition heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1296-1305. doi:10.1037/a0016565

Hilbig, B. E., Pohl, R. F., & Bröder, A. (2009). Criterion knowledge: A moderator of using the recognition heuristic? *Journal of Behavioral Decision Making, 22,* 510-522. doi:10.1002/bdm.644

Hilbig, B. E., & Richter, T. (2011). Homo heuristicus outnumbered: Comment on Gigerenzer and Brighton (2009). *Topics in Cognitive Science, 3*, 187-196. doi:10.1111/j.1756-8765.2010.01123.x

Hilbig, B. E., Scholl, S. G., & Pohl, R. F. (2010). Think or blink – is the recognition heuristic an "intuitive" strategy? *Judgment and Decision Making*, *5*, 300-309. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.177.840&rep=rep1&type=pdf

Hill, B. D., Foster, J. D., Elliott, E. M., Shelton, J. T., McCain, J., & Gouvier, W. D. (2013). Need for cognition is related to higher general intelligence, fluid intelligence, and crystallized intelligence, but not working memory. *Journal of Research in Personality*, *47*, 22-25. doi:10.1016/j.jrp.2012.11.001

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, *57*, 253-270. doi:10.1037/h0023816

Horn, S. S., Pachur, T., & Mata, R. (2015). How does aging affect recognition-based inference? A hierarchical Bayesian modeling approach. *Acta Psychologica*, *154*, 77-85. doi:10.1016/j.actpsy.2014.11.001

Hoyer, W. D., & Brown, S. P. (1990). Effects of brand awareness on choice for a common, repeat-purchase product. *Journal of Consumer Research*, *17*, 141-148. doi:10.1086/208544

Hu, X., & Batchelder, W. H. (1994). The statistical analysis of general processing tree models with the EM algorithm. *Psychometrika*, *59*, 21-47. doi:10.1007/BF02294263

Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, *56*, 326-338. doi:10.1037/0022-3514.56.3.326

Kane, M. J., Conway, A. R. A., & Hambrick, D. Z. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131,* 66-71. doi:10.1037/0033-2909.131.1.66

Kantner, J., & Lindsay, D. S. (2012). Response bias in recognition memory as a cognitive trait. *Memory & Cognition*, *40*, 1163-1177. doi:10.3758/s13421-012-0226-0

Kantner, J., & Lindsay, D. S. (2014). Cross-situational consistency in recognition memory response bias. *Psychonomic Bulletin & Review*, *21*, 1272-1280. doi:10.3758/s13423-014-0608-3

Kämmer, J. E., Gaissmaier, W., Reimer, T., & Schermuly, C. C. (2014). The adaptive use of recognition in group decision making. *Cognitive Science, 38,* 911-942. doi:10.1111/cogs.12110

Keller, J., Bohner, G., & Erb, H. (2000). Intuitive und heuristische Urteilsbildung – verschiedene Prozesse? Präsentation einer deutschen Fassung des 'Rational-Experiential Inventory' sowie neuer Selbstberichtskalen zur Heuristiknutzung. *Zeitschrift für Sozialpsychologie, 31,* 87-101. doi:10.1024//0044-3514.31.2.87

Klauer, K. C. (2006). Hierarchical multinomial processing tree models: A latent-class approach. *Psychometrika, 71*, 7-31. doi:10.1007/s11336-004-1188-3

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika, 75,* 70-98. doi:10.1007/S11336-009-9141-0

LaBar, K. S., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (1999). Neuroanatomic overlap of working memory and spatial attention networks: A functional MRI comparison within subjects. *NeuroImage*, *10*, 695-704. doi:10.1006/nimg.1999.0503

Larrick, R. P., Nisbett, R. E. & Morgan, J. N. (1993). Who uses the cost-benefit rules of choice? Implications for the normative status of microeconomic theory. *Organizational Behavior and Human Decision Processes, 56,* 331-347. doi:10.1006/obhd.1993.1058

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course.* New York, NY: Cambridge University Press.

Lozano, J. H. (2016). Personality and behavior in social dilemmas: Testing the situational strength hypothesis and the role of hypothetical versus real incentives. *Journal of Personality, 84,* 71-78. doi:10.1111/jopy.12139

Macdonald, E. K., & Sharp, B. M. (2000). Brand awareness effects on consumer decision making for a common, repeat purchase product: A replication. *Journal of Business Research*, *48*, 5-15. doi:10.1016/S0148-2963(98)00070-8

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Mahoney, K. T., Buboltz, W., Levin, I. P., Doverspike, D., & Svyantek, D. J. (2011). Individual differences in a within-subjects risky-choice framing study. *Personality and Individual Differences*, *51*, 248-257. doi:10.1016/j.paid.2010.03.035

Mann, L. (1973). Differences between reflective and impulsive children in tempo and quality of decision making. *Child Development*, *44*, 274-279. doi:10.2307/1128047

Marewski, J. N., Gaissmaier, W., Schooler, L. J., Goldstein, D. G., & Gigerenzer, G. (2010). From recognition to decisions: Extending and testing recognition-based models for multialternative inference. *Psychonomic Bulletin & Review, 17*, 287-309. doi:10.3758/PBR.17.3.287

Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision maker: Cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging*, *22*, 796-810. doi:10.1037/0882-7974.22.4.796

Matzke, D., Dolan, C. V., Batchelder, W. H., & Wagenmakers, E.-J. (2015). Bayesian estimation of multinomial processing tree models with heterogeneity in participants and items. *Psychometrika*, *80,* 205-235. doi:10.1007/s11336-013-9374-9

Matzke, D., Lee, M. D., & Wagenmakers, E.-J. (2013). Multinomial processing trees. In M. D. Lee, & E.-J. Wagenmakers (Eds.), *Bayesian cognitive modeling: A practical course* (pp. 187-195). New York, NY: Cambridge University Press.

Matzke, D., Ly, A., Selker, R., Weeda, W. D., Scheibehenne, B., Lee, M. D., & Wagenmakers, E.-J. (2016). *Bayesian inference for correlations in the presence of estimation uncertainty and measurement error.* Manuscript in preparation.

McCabe, D. P., Roediger, H. L., III, McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, *24*, 222-243. doi:10.1037/a0017619

McCloy, R. A., Beaman, C. P., Frosch, C. A. and Goddard, K. (2010). Fast and frugal framing effects? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1043-1052. doi:10.1037/a0019693

McCloy, R., Beaman, C. P., & Smith, P. T. (2008). The relative success of recognition-based inference in multichoice decisions. *Cognitive Science*, *32*, 1037-1048. doi:10.1080/03640210802152319

Michalkiewicz, M., Arden, K., & Erdfelder, E. (2016). *Do smarter people make better decisions? The influence of intelligence on adaptive use of the recognition heuristic.* Manuscript submitted for publication.

Michalkiewicz, M., Coolin, A., & Erdfelder, E. (2013, August). *Individual differences in use of the recognition heuristic.* Poster presented at the 46th annual meeting of the Society of Mathematical Psychology, Potsdam, Germany.

Michalkiewicz, M., Hilbig, B. E., Erdfelder, E., Keller, J., & Bless, H. (2012, April). *The role of regulatory focus in use of the recognition heuristic.* Paper presented at the 54th Conference of Experimental Psychologists, Mannheim, Germany.

Michalkiewicz, M., & Erdfelder, E. (2016). Individual differences in use of the recognition heuristic are stable across time, choice objects, domains, and presentation formats. *Memory & Cognition*. Advance online publication. doi:10.3758/s13421-015-0567-6

Michalkiewicz, M., Minich, B., & Erdfelder, E. (2016). *Explaining individual differences in fast-and-frugal decision making: The impact of need for cognition and faith in intuition on use of the recognition heuristic.* Manuscript submitted for publication.

Mischel, W. (1968). *Personality and Assessment.* New York, NY: Wiley.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, *80*, 252-283. doi:10.1037/h0035002

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49-100. doi:10.1006/cogp.1999.0734

Mohammed, S., & Schwall, A. (2009). Individual differences and decision making: What we know and where we go from here. In G. P. Hodgkinson, & J. K. Ford (Eds.), *International Review of Industrial and Organizational Psychology* (Volume 24, pp. 249-312). Chichester, UK: Wiley-Blackwell.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., … Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77-101. doi:10.1037/0003-066X.51.2.77

Newell, B. R., & Fernandez, D. (2006). On the binary quality of recognition and the inconsequentiality of further knowledge: Two critical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19*, 333-346. doi:10.1002/bdm.531

Newell, B. R., & Shanks, D. R. (2004). On the role of recognition in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 923-935. doi:10.1037/0278-7393.30.4.923

Odum, A. L. (2011). Delay discounting: Trait variable? *Behavioral Processes, 87*, 1-9. doi:10.1016/j.beproc.2011.02.007

Oeusoonthornwattana, O., & Shanks, D. R. (2010). I like what I know: Is recognition a non-compensatory determiner of consumer choice. *Judgment and Decision Making*, *5*, 310-325. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/ download?doi=10.1.1.177.329&rep=rep1&type=pdf

Oppenheimer, D. M. (2003). Not so fast! (and not so frugal!): Rethinking the recognition heuristic. *Cognition, 90,* B1-B9. doi:10.1016/S0010-0277(03)00141-0

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology, 76,* 972-987. doi:10.1037/0022-3514.76.6.972

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 534-552. doi:10.1037/0278-7393.14.3.534

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York, NY: Cambridge University Press.

Pachur, T. (2011). The limited value of precise tests of the recognition heuristic. *Judgment and Decision Making, 6*, 413-422. Retrieved from http://search.proquest.com/docview/1011294684?accountid=14570

Pachur, T., & Biele, G. (2007). Forecasting from ignorance: The use and usefulness of recognition in lay predictions of sports events. *Acta Psychologica, 125*, 99-116. doi:10.1016/j.actpsy.2006.07.002

Pachur, T., Bröder, A., & Marewski, J. N. (2008). The recognition heuristic in memory-based inference: Is recognition a non-compensatory cue? *Journal of Behavioral Decision Making, 21*, 183-210. doi:10.1002/bdm.581

Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 983-1002. doi:10.1037/0278-7393.32.5.983

Pachur, T., & Marinello, G. (2013). Expert intuitions: How to model the decision strategies of airport customs officers? *Acta Psychologica, 144*, 97-103. doi:10.1016/j.actpsy.2013.05.003

Pachur, T., Mata, R., & Schooler, L. J. (2009). Cognitive aging and the adaptive use of recognition in decision making. *Psychology and Aging, 24*, 901-915. doi:10.1037/a0017211

Pachur, T., & Spaar, M. (2015). Domain-specific preferences for intuition and deliberation in decision making. *Journal of Applied Research in Memory and Cognition, 4,* 303-311. doi:10.1016/j.jarmac.2015.07.006

Pachur, T., Todd, P. M., Gigerenzer, G., Schooler, L. J., & Goldstein, D. G. (2011). The recognition heuristic: A review of theory and tests. *Frontiers in Psychology, 2,* 1-14. doi:10.3389/fpsyg.2011.00147

Phillips, W. J., Fletcher, J. M., Marks, A. D. G., & Hine, D. W. (2016). Thinking styles and decision making: A meta-analysis. *Psychological Bulletin, 142,* 260-290. doi:10.1037/bul0000027

Pohl, R. F. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19*, 251-271. doi:10.1002/bdm.522

Pohl, R. F. (2011). On the use of recognition in inferential decision making: An overview of the debate. *Judgment and Decision Making, 6,* 423-438. Retrieved from http://journal.sjdm.org/11/rh19/rh19.pdf

Pohl, R. F., Erdfelder, E., Hilbig, B. E., Liebke, L., & Stahlberg, D. (2013). Effort reduction after self-control depletion: The role of cognitive resources in use of simple heuristics. *Journal of Cognitive Psychology, 25,* 267-276. doi:10.1080/20445911.2012.758101

Pohl, R. F., von Massow, F., & Beckmann, B. (2016). *Developmental differences in using an ecologically valid decision strategy: The case of the recognition heuristic.* Manuscript submitted for publication.

Reimer, T., & Hoffrage, U. (2003). Information aggregation in groups: The approach of simple group heuristics (SIGH). In R. Alterman & D. Kirsch (Eds.), *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society* (pp. 982–987). Boston, MA: Cognitive Science Society.

Reimer, T., & Hoffrage, U. (2005). Can simple group heuristics detect hidden profiles in randomly generated environments*? Swiss Journal of Psychology, 64,* 21-37. doi:10.1024/1421-0185.64.1.21

Reimer, T., & Hoffrage, U. (2006). The ecological rationality of simple group heuristics: Effects of group member strategies on decision accuracy. *Theory and Decision, 60,* 403-438. doi:10.1007/s11238-005-4750-2

Reimer, T., & Katsikopoulos, K. V. (2004). The use of recognition in group decision making. *Cognitive Science, 28,* 1009-1029. doi:10.1016/j.cogsci.2004.06.004

Revelle, W. (1987). Personality and motivation: Sources of inefficiency in cognitive performance. *Journal of Research in Personality*, *21*, 436-452. doi:10.1016/0092-6566(87)90031-6

Richter, T., & Späth, P. (2006). Recognition is used as one cue among others in judgment and decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 150-162. doi:10.1037/0278-7393.32.1.150

Riefer, D. M., & Batchelder, W. H. (1991). Statistical inference for multinomial processing tree models. In J.-P. Doignon & J.-C. Falmagne (Eds.),

Mathematical psychology: Current developments (pp. 313-335). New York, NY: Springer.

Rindermann, H., & Neubauer, A. C. (2001). The influence of personality on three aspects of cognitive performance: Processing speed, intelligence and school performance. *Personality and Individual Differences*, *30*, 829-842. doi:10.1016/S0191-8869(00)00076-3

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*, 638-641. doi:10.1037/0033-2909.86.3.638

Scheibehenne, B., & Bröder, A. (2007). Predicting Wimbledon tennis results 2005 by mere player name recognition. *International Journal of Forecasting, 23,* 415-426. doi:10.1016/j.ijforecast.2007.05.006

Scheibehenne, B., & Pachur, T. (2015). Using Bayesian hierarchical parameter estimation to assess the generalizability of cognitive models of choice. *Psychonomic Bulletin & Review*, *22*, 391-407. doi:10.3758/s13423-014-0684-4

Serwe, S., & Frings, C. (2006). Who will win Wimbledon? The recognition heuristic in predicting sports events. *Journal of Behavioral Decision Making, 19*, 321-332. doi:10.1002/bdm.530

Shiloh, S., Koren, S., & Zakay, D. (2001). Individual differences in compensatory decision-making style and need for closure as correlates of subjective decision complexity and difficulty. *Personality and Individual Differences, 30,* 699-710. doi:10.1016/S0191-8869(00)00073-8

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129-138. doi:10.1037/h0042769201

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41,* 1-19. doi:10.1146/annurev.ps.41.020190.000245

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3–22. doi:10.1037/0033-2909.119.1.3

Smith, J. B., & Batchelder, W. H. (2008). Assessing individual differences in categorical data. *Psychonomic Bulletin & Review, 15*, 713-731. doi:10.3758/PBR.15.4.713

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology, 15,* 72-101. doi:10.2307/1412159

Stahl, C., & Degner, J. (2007). Assessing automatic activation of valence: A multinomial model of EAST performance. *Experimental Psychology*, *54*, 99-112. doi:10.1027/1618-3169.54.2.99

Stahl, C., & Klauer, K. C. (2007). HMMTree: A computer program for latent-class hierarchical multinomial processing tree models. *Behavior Research Methods, 39*, 267-273. doi:10.3758/BF03193157

Stahl, C., & Klauer, K. C. (2008). A simplified conjoint recognition paradigm for the measurement of gist and verbatim memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 570-586. doi:10.1037/0278-7393.34.3.570

Stahl, C., & Klauer, K. C. (2009). Measuring phantom recollection in the simplified conjoint recognition paradigm. *Journal of Memory and Language*, *60*, 180-193. doi:10.1016/j.jml.2008.08.001

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127,* 161-188. doi:10.1037/0096-3445.127.2.161

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23,* 645-726. doi:10.1017/S0140525X00003435

Sternberg, R. J., & Salter, W. (1982). Conceptions of intelligence. In R.J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 3- 29). Cambridge, MA: Cambridge University Press.

Strobl, C., Wickelmaier, F., & Zeileis, A. (2011). Accounting for individual differences in Bradley-Terry models by means of recursive partitioning. *Journal of Educational and Behavioral Statistics*, 36, 135-153. doi:10.3102/1076998609359791

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning. The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293-315. doi:10.1037/0033-295X.90.4.293

Volz, K. G., Schooler, L. J., Schubotz, R. I., Raab, M., Gigerenzer, G., & von Cramon, D. Y. (2006). Why you think Milan is larger than Modena: Neural correlates of the recognition heuristic. *Journal of Cognitive Neuroscience*, *18*, 1924-1936. doi:10.1162/jocn.2006.18.11.1924

von Stumm, S. (2013). Investment traits and intelligence in adulthood. Assessment and associations. *Journal of Individual Differences, 34,* 82-89. doi:10.1027/1614-0001/a000101

Wickelmaier, F., & Zeileis, A. (2013, August). *Investigating the effects of subject covariates in multinomial processing tree models by recursive partitioning.* Paper presented at the 46th Annual Meeting of the Society of Mathematical Psychology, Potsdam, Germany.

Witkin, H. A., Goodenough, D. R., & Karp, S. A. (1967). Stability of cognitive style from childhood to young adulthood. *Journal of Personality and Social Psychology*, *7*, 291-300. doi:10.1037/h0025070

Yechiam, E., & Busemeyer, J. R. (2008). Evaluating generalizability and parameter consistency in learning models. *Games and Economic Behavior*, *63*, 370-394. doi:10.1016/j.geb.2007.08.011

Zimmerman, I., Redker, C., & Gibson, B. (2011). The role of faith in intuition, need for cognition and method of attitude formation in implicit-explicit brand attitude relationship strength. *Journal of Consumer Psychology*, *21*, 290-301. doi:10.1016/j.jcps.2011.04.001

Individual differences in use of the RH

Statement of Originality

I hereby declare that I am the sole author of this thesis and have made use of no other sources than those cited in this work.

Mannheim, March 2016   Martha Michalkiewicz

Co-Authors' Statements

Co-author's statement (E. Erdfelder)

I hereby confirm that the following articles were primarily conceived and written by Dipl.-Math. Martha Michalkiewicz, School of Social Sciences, University of Mannheim.

Michalkiewicz, M., & Erdfelder, E. (2016). Individual differences in use of the recognition heuristic are stable across time, choice objects, domains, and presentation formats. *Memory & Cognition, 44,* 454-468. doi:10.3758/s13421-015-0567-6

Michalkiewicz, M., Arden, K., & Erdfelder, E. (2016). *Do smarter people make better decisions? The influence of intelligence on adaptive use of the recognition heuristic.* Manuscript submitted for publication.

Michalkiewicz, M., Minich, B., & Erdfelder, E. (2016). *Explaining individual differences in fast-and-frugal decision making: The impact of need for cognition and faith in intuition on use of the recognition heuristic.* Manuscript submitted for publication.

Mannheim, March 2016   Prof. Dr. Edgar Erdfelder

Co-author's statement (K. Arden)

I hereby confirm that the following article was primarily conceived and written by Dipl.-Math. Martha Michalkiewicz, School of Social Sciences, University of Mannheim.

Michalkiewicz, M., Arden, K., & Erdfelder, E. (2016). *Do smarter people make better decisions? The influence of intelligence on adaptive use of the recognition heuristic.* Manuscript submitted for publication.

Mannheim, March 2016   Katja Arden, B. Sc.

Co-author's statement (B. Minich)

I hereby confirm that the following article was primarily conceived and written by Dipl.-Math. Martha Michalkiewicz, School of Social Sciences, University of Mannheim.

Michalkiewicz, M., Minich, B., & Erdfelder, E. (2016). *Explaining individual differences in fast-and-frugal decision making: The impact of need for cognition and faith in intuition on use of the recognition heuristic.* Manuscript submitted for publication.

Mannheim, March 2016   Barbara Minich, B. Sc.

Appendix: Copies of Manuscripts

Manuscript 1

Michalkiewicz, M., & Erdfelder, E. (2016). Individual differences in use of the
recognition heuristic are stable across time, choice objects, domains, and
presentation formats. *Memory & Cognition, 44,* 454-468. doi:10.3758/s13421-
015-0567-6

Manuscript 2

Michalkiewicz, M., Arden, K., & Erdfelder, E. (2016). *Do smarter people make better
decisions? The influence of intelligence on adaptive use of the recognition
heuristic.* Manuscript submitted for publication.

Manuscript 3

Michalkiewicz, M., Minich, B., & Erdfelder, E. (2016). *Explaining individual
differences in fast-and-frugal decision making: The impact of need for cognition
and faith in intuition on use of the recognition heuristic.* Manuscript submitted
for publication.

CrossMark

Individual differences in use of the recognition heuristic are stable across time, choice objects, domains, and presentation formats

Martha Michalkiewicz[1] · Edgar Erdfelder[1]

Abstract The recognition heuristic (RH) is a simple decision strategy that performs surprisingly well in many domains. According to the RH, people decide on the basis of recognition alone and ignore further knowledge when faced with a recognized and an unrecognized choice object. Previous research has revealed noteworthy individual differences in RH use, suggesting that people have preferences for using versus avoiding this strategy that might be causally linked to cognitive or personality traits. However, trying to explain differences in RH use in terms of traits presupposes temporal and cross-situational stability in use of the RH, an important prerequisite that has not been scrutinized so far. In a series of four experiments, we therefore assessed the stability in RH use across (1) time, (2) choice objects, (3) domains, and (4) presentation formats of the choice objects. In Experiment 1, participants worked on the same inference task and choice objects twice, separated by a delay of either one day or one week. Experiment 2 replicated Experiment 1 using two different object sets from the same domain, whereas Experiment 3 assessed the stability of RH use across two different domains. Finally, in Experiment 4 we investigated stability across verbal and pictorial presentation formats of the choice objects. For all measures of RH use proposed so far, we found strong

evidence for both temporal and cross-situational stability in use of the RH. Thus, RH use at least partly reflects a person-specific style of decision making whose determinants await further research.

Which city is more populous: Tokyo or Busan? If you recognize Tokyo but not Busan, you can use a simple inference strategy: the fast-and-frugal recognition heuristic (RH; Goldstein & Gigerenzer, 2002). According to the RH, a person should choose the recognized object and ignore any further knowledge. Thus, when following the RH, you would choose Tokyo simply because you recognize it. Alternatively, you can deliberately integrate knowledge available over and above recognition—for instance, that Tokyo has an international airport and that cities with an international airport are (most often) more populous. In this case, you would arrive at the same conclusion with both decision strategies. However, which factors are responsible for using the RH versus integrating further knowledge?

There is a large body of research on the situational determinants of RH use. In general, RH use increases, the greater the importance of a quick decision (Hilbig, Erdfelder, & Pohl, 2012; Pachur & Hertwig, 2006) and the higher the validity of the recognition cue (Castela, Kellen, Erdfelder, & Hilbig, 2014; Hilbig, Erdfelder, & Pohl, 2010; Pachur, Mata, & Schooler, 2009; Pohl, 2006; Scheibehenne & Bröder, 2007). By contrast, integration of further knowledge increases as knowledge becomes more easily available and easier to integrate (Bröder & Eichler, 2006; Glöckner & Bröder, 2011; Hilbig, Michalkiewicz, Castela, Pohl, & Erdfelder, 2015;

✉ Martha Michalkiewicz
michalkiewicz@psychologie.uni-mannheim.de

[1] Department of Psychology, School of Social Sciences, University of Mannheim, Schloss, Ehrenhof-Ost, 68131 Mannheim, Germany

🜨 Springer

Hilbig, Pohl, & Bröder, 2009; Newell & Fernandez, 2006; Richter & Späth, 2006). In sum, it is quite well established that participants adjust their RH use according to situational factors (for reviews, see Gigerenzer & Goldstein, 2011; Pachur, Todd, Gigerenzer, Schooler, & Goldstein, 2011; Pohl, 2011).

However, studies that have addressed situational factors have also revealed large individual differences in RH use (Hilbig & Richter, 2011; Marewski, Gaissmaier, Schooler, Goldstein, & Gigerenzer, 2010; Newell & Shanks, 2004; Pachur, Bröder, & Marewski, 2008). As was argued by Gigerenzer and Brighton (2009, p. 133), "in virtually every task we find individual differences in strategies."

Figure 1 displays the individual proportions of RH use on the basis of data from Hilbig and Pohl (2009, Exp. 1), assessed with the *r*-model (Hilbig et al., 2010) that we will describe in detail below. Why would RH use differ to such a degree between participants under constant context conditions? As is indicated by the standard errors illustrated in Fig. 1, the observed heterogeneity is too large to be attributable to error variance only. In fact, a goodness-of-fit test of the homogeneity hypothesis that all 24 individual proportions of RH use are equal reveals a clear misfit [$\Delta G^2(23) = 72.8$, $p < .001$]. This suggests that the heterogeneity might reflect individual preferences for certain strategies that are not determined by the context. This, in turn, gives rise to the question: Do individual traits underlie RH use?

Indeed, there is some evidence that different groups of people prefer different strategies. In particular, Pachur et al. (2009) showed that elderly people use the RH more often than young adults do (see also Horn, Pachur, & Mata, 2015). Extending this line of research to the life span, Pohl, von Massow, and Beckmann (2015) detected a nonmonotonic trend in RH use in younger age groups: Preadolescent children and young adults used the RH about equally often, whereas adolescents used it more frequently. Moreover,
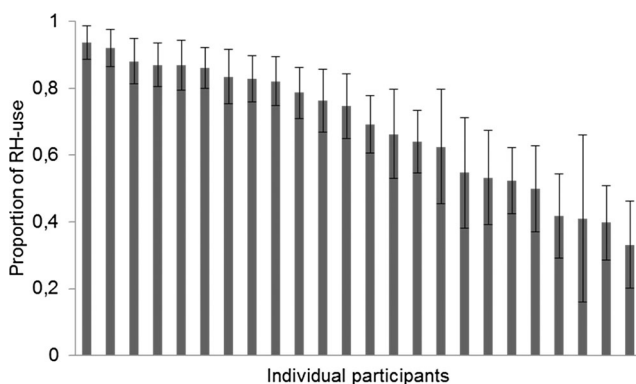
exploring a different source of individual differences, Hilbig and Pohl (2008) found that more knowledgeable people tend to rely less on the RH. These examples show that groups of individuals may differ significantly from each other in RH use even if the decision context is kept constant. In addition, at least one study successfully examined the relationship between RH use and personality traits: Hilbig (2008) demonstrated that neuroticism was positively related to RH use. Results like this one encourage a search for traits as sources of individual differences in RH use.

However, considerable evidence also shows how difficult it is to find associations between strategy use and individual traits. For instance, apart from neuroticism, Hilbig (2008) also investigated agreeableness, conscientiousness, openness, and extraversion, but did not find any substantial effect on RH use. Similarly, Pachur et al. (2009) tested associations between measures of inhibitory control and RH use without finding evidence for substantial correlations. Furthermore, Bröder (2012) summarized multiple studies on the take-the-best heuristic (TTB; Gigerenzer & Goldstein, 1999) and various cognitive and personality traits. Only intelligence was found to affect adaptive use of the TTB heuristic, depending on the environmental payoff structure (see also Bröder, 2003). Notably, none of the remaining variables covered by Bröder's (2012) review, including need for cognition, impulsivity, and the Big Five personality traits, showed any substantial relation.

Given the equivocal evidence on traits as determinants of decision strategy use, we argue that prior to trying to explain variability in RH use in terms of cognitive or personality traits, an important precondition should be checked: stability in use of the RH. The need for stability assessment has previously been pointed out by Bröder and Newell (2008, p. 208): "It is yet an open question whether the different strategy preferences diagnosed in a one-shot assessment of an experiment will turn out to be stable across tasks and situations." If RH use turns out to be stable, it makes sense to search for relations between RH use and individual traits. If, in contrast, stability cannot be shown—that is, if RH use varies haphazardly within individuals across situations, then it will hardly be possible to find replicable relations between RH use and cognitive or personality traits. Of course, stability does not imply that each individual behaves identically in all situations—that is, exhibits exactly the same level of RH use everywhere. Rather, it means that individual behavior is "meaningfully consistent" (Roberts, 2009, p. 139)—for example, that participants showing higher than average RH use in one decision context will also tend to show higher than average RH use in a second context.

Notably, several studies have already emphasized stability in behaviors related to judgment and decision making. Thus, evidence suggesting the stability of RH use would fit nicely into related lines of research. For instance, Witkin,

Fig. 1 Individual proportions of RH use per participant for the data from Hilbig and Pohl (2009, Exp. 1). RH use is estimated via the *r* parameter of the *r*-model (Hilbig et al., 2010) using multiTree (Moshagen, 2010) and ordered by size. Error bars illustrate standard errors.

Goodenough, and Karp (1967) showed that the ability to ignore the visual context in perceptual judgments—known as *field independence*—is a stable cognitive style from childhood to young adulthood, at least. Furthermore, Couch and Keniston (1960) investigated acquiescence bias in terms of consistency over time and generality over tests. More recently, Kantner and Lindsay (2012, 2014) found response bias in recognition tasks to be stable across time, stimulus materials, and item presentation formats. Similarly, Aminoff and colleagues (2012) demonstrated stability in criterion shifting in recognition memory across different presentation formats. Furthermore, Odum (2011) reanalyzed prior studies on delay discounting, and the results suggested stability across time for up to one year using many different stimulus materials.

Recently, the focus has been shifted to systematic tests of parameter stability of cognitive models. For instance, Yechiam and Busemeyer (2008) evaluated the stability of several learning models for repeated choice problems across different tasks. Also, Glöckner and Pachur (2012) examined parameter stability with respect to cumulative prospect theory across time (see also Scheibehenne & Pachur, 2015). These are just some examples of stability research, and many more could be listed.

Following a route similar to those in the studies outlined above, we will analyze four different aspects of stability—namely, stability across (1) time, (2) choice objects, (3) domains, and (4) presentation formats. For this purpose, we conducted four experiments in which participants completed two sets of inference tasks. Stability was measured as the test–retest correlation between RH use in Tests 1 and 2. In Experiment 1, we assessed stability across time for delays of one day versus one week, using exactly the same choice objects in both tests. Experiment 2 was designed to replicate Experiment 1. This time, however, different choice objects were drawn from the same domain on the two tests, providing for an assessment of stability across disjoint sets of objects. To further analyze the influence of the choice materials, we conducted Experiment 3, in which we assessed the stability of RH use across different domains. Finally, to examine stability across presentation formats, we designed Experiment 4, in which we used names versus pictures as formats of object presentation.

Note that, in all four experiments, we specifically opted to investigate stability factors that are unconfounded with the overall level of RH use. Investigating the same people twice—even when using different materials or presentation formats—is usually assumed not to affect the overall level of RH use, provided that the recognition and knowledge validities do not differ between tests. To the degree that we succeeded in implementing these conditions, we expected neither main effects of situational factors nor interaction effects with individual factors on RH use, enabling us to assess the influence of individual differences without confounds.

General method

When investigating stability in use of the RH, it is important to ensure that both strategies under consideration—the RH and knowledge use—are applicable in the current context. For this purpose, several conditions must be fulfilled. First, participants must recognize at least one, and at most all but one, of the objects in order to apply the RH in the first place. Optimally, participants should recognize half of the objects to maximize the proportion of cases in which the RH can be applied. Second, participants must obviously have some kind of knowledge about the set of objects and the question of interest. If nothing but recognition information is available, participants obviously cannot apply more elaborated strategies incorporating further knowledge. Third, the validity of the recognition cue α (i.e., the proportion of cases in which choosing the recognized object leads to a correct response) and the validity of knowledge β (i.e., the proportion of cases in which the application of knowledge leads to a correct response) should both be greater than chance. These conditions render both the RH and the use of knowledge reasonable strategies. When selecting the materials for our experiments, we aimed at satisfying all of these requirements.

To analyze our data, we primarily relied on the *r*-model (Fig. 2; Hilbig et al., 2010), a multinomial processing tree model (Batchelder & Riefer, 1999; Erdfelder et al., 2009) tailored to measure RH use, as defined by Goldstein and Gigerenzer (2002). Specifically, if exactly one object is recognized, participants will either apply the RH with probability r or make use of further knowledge with probability $1 - r$. We focused on this model because it successfully decontaminates the probability of RH use from the effects of knowledge-based strategies that might also lead to choice of the recognized object (see Hilbig, 2010).

In the present analyses, we applied the latent-trait approach to multinomial processing tree models (Klauer, 2010) because it elegantly handles variability in parameters between individuals. For this purpose, we constructed a hierarchical version of the *r*-model (Fig. 3) based on the implementation by Matzke, Dolan, Batchelder, and Wagenmakers (2015) and extended it to account for the data of two test occasions simultaneously. Compared to standard correlational analyses, the latent-trait approach has one main advantage (Klauer, 2010; Matzke et al., 2015): It allows for the joint estimation of model parameters and the correlations between parameters in a single step. The estimated correlations are thus automatically adjusted for the uncertainty in the individual parameter estimates; that is, the model estimates the correlation of the true scores decontaminated from error influences. For a comprehensive introduction to hierarchical models and their advantages, see, for instance, Lee and Wagenmakers (2013).

Extension of the single-test hierarchical *r*-model to a two-test version is straightforward. We estimated the parameters of this extended model within the Bayesian framework using Markov chain Monte Carlo sampling employing OpenBUGS
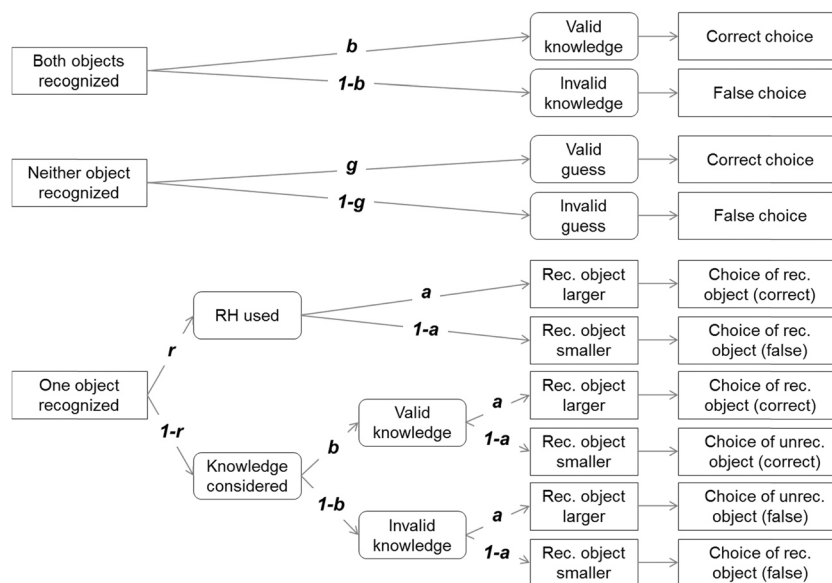
Fig. 2 Illustration of the *r*-model. Displayed are the three object pairs that can occur in a paired-comparison task: knowledge pairs (both objects recognized), guessing pairs (neither object recognized), and recognition pairs (exactly one object recognized). Participants' responses are assigned to eight categories and accounted for by four latent parameters—namely, recognition validity (parameter *a*), knowledge validity (parameter *b*), the probability of correct guessing (parameter *g*), and most importantly, the probability of RH use (parameter *r*). Adapted from "One-Reason Decision Making Unveiled: A Measurement Model of the Recognition Heuristic," by Hilbig et al. 2010, *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, p. 125. Copyright 2010 by the American Psychological Association.

(Lunn, Spiegelhalter, Thomas, & Best, 2009) through R2WinBUGS (Sturtz, Ligges, & Gelman, 2005).[1] For each analysis, we ran three chains with 500,000 iterations each, using a thinning rate of 10, and discarded the first 100,000 iterations as a burn-in period. Chain convergence was reached for all estimated parameters ($R < 1.01$; Gelman, Carlin, Stern, & Rubin, 2004). Also, the effective sample sizes were sufficient to trust the parameter estimates (Kruschke, 2014).[2] As is common practice in Bayesian analysis, we will report the means of the posterior distributions together with their 95 % Bayesian credible intervals (BCI). In particular, μ^{r1} and μ^{r2} are interpreted as the group-level estimates of RH use on Test Occasions 1 and 2, respectively. On the basis of the covariance matrix Σ, the standard deviations σ^{r1} and σ^{r2} as well as the correlations between parameters $\rho_{r1,r2}$ are derived. In this context, standard deviations reflect the variation between participants, being close to zero when participants are rather homogeneous and large when there are substantial individual differences.

Experiment 1

To assess stability over time, we tested whether a given participant would show similar levels of RH use for a set of choice objects on two different points in time.

Method

Design and procedure To study the RH, we employed the frequently used city-size task (Goldstein & Gigerenzer, 2002; Hilbig & Pohl, 2009; Pachur et al., 2008) including, in random order, a paired-comparison task and a recognition task. In the comparison task, participants were asked to decide for pairs of cities which of the two cities was more populous. In the recognition task, participants had to indicate for all cities whether or not they had heard of the city before the experiment.

To test our hypothesis, participants were randomly assigned to one of two groups and worked on the city-size task twice, separated by a delay of either one day (*day group*) or one week (*week group*). To render the procedures equivalent for both groups, all participants completed three sessions—the initial session and the two following sessions one day and one week later—at exactly the same time of day. In Session 1, all participants first worked on the city-size task and then completed two unrelated experiments to render the intention of the study less obvious. In Sessions 2 and 3, depending on the group, participants either worked on the city-size task for the second time or again completed an unrelated experiment. To control for contamination of the results, we asked participants two questions after they had completed the city-size task for the second time: (1) whether they had tried to memorize their responses in the first session, and (2) whether they had looked up city sizes after the first session. Participants received course credit or a flat fee of €10 after (and only after) completion of all three sessions. We refrained from using performance-contingent payment here, as

---

[1] The R code, the model file, and a sample data set are provided in the online supplemental materials.
[2] To meet these criteria, we increased the number of iterations per chain to 1 million for the week group of Experiments 1 and 2 as well as for the different group of Experiment 3.
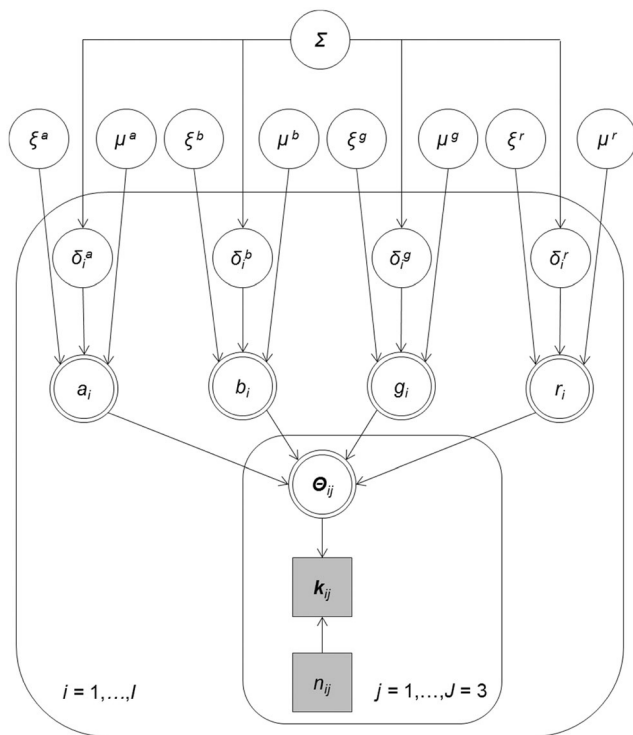
Fig. 3 Illustration of the single-test hierarchical latent-trait r-model (cf. Matzke, Lee, & Wagenmakers, 2013). Displayed are the relations between data and latent model parameters. Observable and unobservable variables are presented as shaded versus unshaded nodes, respectively; discrete and continuous variables are presented as squares versus circular nodes. To-be-estimated and derived variables (which can be parameterized by means of the remaining model parameters) are presented as single-bordered versus double-bordered nodes. The plates indicate replications over I individuals for the $J = 3$ object cases (knowledge, guessing, and recognition cases). For each individual i and each object case j, the vector of category counts $\mathbf{k}_{ij}$ follows a multinomial distribution with probability vector Θ_{ij} and number of observations n_{ij}, as is presented in the r-model (see Fig. 2). The individual model parameters s_i ($s \in \{a, b, g, r\}$) are modeled in a probit-transformed space, $s_i \leftarrow \Phi(\mu^s + \xi^s * \delta_i^s)$, as linear combinations of the group-level mean $\mu^s \sim N(0, 1)$, the multiplicative scale parameter $\xi^s \sim U(0, 100)$, and the individual displacement parameter δ_i^s, drawn from a common multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix Σ, with Σ^{-1} following a Wishart($\mathbf{I}$, 5) distribution.

is often done in this kind of task, because we were concerned that this would encourage participants to look up the city sizes between the first and second tests of the RH.

Material We used exactly the same objects in both repetitions of the city-size task. Specifically, we selected a random sample of 100 cities from the 150 most populous US cities for the recognition tasks. From these items, we randomly created one sample of 300 pairs for the comparison tasks, ensuring that (1) each city appeared exactly six times and (2) recognition and knowledge validity were adequate to render both RH use and knowledge use reasonable strategies to solve the task. To achieve this, we selected the materials on

the basis of the data of pilot experiments on RH use conducted in our lab[3].

Participants A total of 70 student participants were recruited via posters and mailing lists at the University of Mannheim. Six participants dropped out of the experiment by not coming back to the second or third session. The remaining 64 participants consisted of 45 women and 19 men, between 18 and 30 years of age ($M = 21.8$ years, $SD = 2.7$). All participants were native speakers or fluent in German.

Results and discussion

The recognition proportions, recognition validities, and knowledge validities did not differ significantly between groups [all $ts(62) < 1.67$, $ps > .10$, Bayes factors (BF_{10}s) < 0.82].[4] Therefore, the analyses related to these variables are reported for both groups combined. As expected, the mean recognition validity[5] $\overline{\alpha}$ and the mean knowledge validity[6] $\overline{\beta}$ showed that both strategies—the RH and knowledge use—were clearly better than guessing ($\overline{\alpha} = .77$, $SD = .06$; $\overline{\beta} = .67$, $SD = .08$, and $\overline{\alpha} = .77$, $SD = .06$; $\overline{\beta} = .66$, $SD = .08$, for Tests 1 and 2, respectively) [all $ts(63) > 15.0$, $ps < .001$, BF_{10}s $> 1,000$]. Hence, the application of either of these strategies was reasonable. Furthermore, participants in the initial session recognized about half of the cities ($M = 49.1$ objects, $SD = 11.4$), resulting in a sufficient number of recognition cases. Surprisingly, participants recognized more cities on the second test of the RH ($M = 58.6$ objects, $SD = 19.8$) than on the first [$t(63) = 4.71$, $p < .001$, $BF_{10} > 1,000$]. This was most probably due to confusion of real recognition (i.e., cities seen before the experiment) and familiarity induced by the presentation of the cities in the initial session. Hence, since the recognition judgments collected on the second test were obviously biased, we based all measures on the recognition judgments obtained in the first test only. However, using the original recognition judgments of each session did not change the results substantially (see Table 2 in the Appendix).

To further control for possible confounds, we analyzed the control questions. All participants confirmed that they had not looked up the city sizes. This was validated by the numbers of correct answers in the comparison tasks across the two tests:

---

[3] The stimulus materials of all experiments are provided in the online supplemental materials.
[4] Bayes factors were computed using the BayesFactor R package (Rouder, Speckman, Sun, Morey, & Iverson, 2009) and interpreted following the classification by Jeffreys (1961).
[5] Recognition validity is computed as the proportion of recognition pairs where the recognized object represents the correct choice. It matches the estimates for parameter a of the hierarchical r-model.
[6] Knowledge validity is computed as the proportion of correct choices in knowledge pairs. It matches the estimates for parameter b of the hierarchical r-model.

Indeed, we even found a slight decrease ($M = 201.9$ answers, $SD = 12.5$, and $M = 199.0$ answers, $SD = 13.1$, for Tests 1 and 2, respectively) [$t(63) = 2.66$, $p = .01$, $BF_{10} = 3.48$]. Moreover, only nine participants stated that they had tried to memorize their answers during the initial session. Obviously, memorizing the answers to 300 questions, given only 1,351 ms per question (i.e., the average response time across participants and items in the comparison task), and retaining them for up to one week is very unlikely. The actual answers confirmed this: Participants decided differently on the second test than on the first for up to half of the trials ($M = 69.0$, $SD = 18.8$, Min $= 38$, Max $= 142$). In sum, there was no indication that judgments in the second comparison task were biased.

As expected, we found strong heterogeneity in RH use between participants ($\mu^{r1} = .77$ [.74, .80], $\sigma^{r1} = .38$ [.29, .48], and $\mu^{r2} = .70$ [.66, .74], $\sigma^{r2} = .42$ [.33, .53], for Tests 1 and 2, respectively). Unexpectedly, we observed a small decrease in the r parameter between the two test occasions ($\Delta_r = .07$ [.04, .10]), showing that participants used the RH slightly less often in the second test than in the first. However, this difference should not have influenced the core results, because the correlation between RH use on Tests 1 and 2 is independent of the mean level of RH use.

To test the main hypothesis, we examined the correlations between the r parameter on Tests 1 and 2 for both groups separately. Overall, we observed strong positive correlations for both groups ($\rho_{r1,r2} = .80$ [.56, .94], and $\rho_{r1,r2} = .71$ [.39, .91], for the day and week groups, respectively). The small drop in correlations after an extended delay of one week was not reliable, as can be seen by the overlapping BCIs. This provided evidence in favor of our hypothesis that people used the RH consistently across time.

To establish a benchmark against which to compare the sizes of the correlations, we estimated the *within-test correlation* of the r parameter for single tests. More precisely, we split the data of each participant into the first and the last 150 trials. We then estimated the correlations between the parameters of these two parts using the hierarchical r-model. The magnitude of consistency across tests was similar to that observed within a single test ($\rho_{r1,r2} = .73$ [.46, .91] on Test 1 and $\rho_{r1,r2} = .72$ [.47, .89] on Test 2), showing that the delay between task repetitions (0 h, 24 h, or 168 h) had little effect on the test–retest correlation. In sum, the results reflect stability across time up to one week, at least when using the same choice objects repeatedly.

A possible objection against Experiment 1 is that participants perhaps just behaved very similarly when working on exactly the same task and choice objects twice. Why should they change their judgments when facing the same choice objects for the second time, perhaps even remembering (some of) the choices they had made previously? To test whether stability was caused by the invariance of materials only, we conducted a second experiment using different objects in the two inference tasks.

Experiment 2

Method

Design and procedure The design and procedure were identical to those of Experiment 1. Participants were again randomly assigned to one of two groups and worked on the city-size task twice: in the initial session and either one day (*day group*) or one week (*week group*) after the initial session.

Material This time, two disjoint samples of 25 cities were randomly drawn from the 61 most populous world cities for the two recognition tasks. Each of the 25 cities was exhaustively paired, resulting in two samples of 300 pairs for the two comparison tasks. Because we wanted the two iterations of the city-size task to closely resemble each other except for object identity, we selected the materials on the basis of the data of prior experiments (Hilbig et al., 2010, 2012). Thereby, we made sure that both city samples had similar proportions of recognized objects, recognition, and knowledge validities.

Participants A total of 94 student participants were recruited via posters and mailing lists at the University of Mannheim. Six participants completed the first session only, and thus dropped out of the experiment. This resulted in 88 participants, consisting of 49 women and 39 men, between 18 and 59 years of age ($M = 22.7$ years, $SD = 5.3$). All participants were native speakers or fluent in German.

Results and discussion

Five participants had to be excluded from the analyses because they recognized all of the objects or all but one. Because group differences were again negligible [all $ts(81) < 1.21$, $ps > .23$, $BF_{10}s < 0.43$], the analyses of recognition rates, recognition validities, and knowledge validities were based on the data of both groups combined. First, participants recognized on average 17.3 cities ($SD = 2.8$) of Set 1 and 16.5 cities ($SD = 1.8$) of Set 2, resulting in sufficient recognition cases for both sets. Second, the mean recognition and knowledge validities showed that both strategies—the RH and further knowledge—were appropriate—that is, better than guessing ($\overline{\alpha} = .69$, $SD = .08$; $\overline{\beta} = .57$, $SD = .08$, and $\overline{\alpha} = .68$, $SD = .07$; $\overline{\beta} = .61$, $SD = .08$, for Sets 1 and 2, respectively) [all $ts(82) > 7.61$, $ps < .001$, $BF_{10}s > 1,000$]. In sum, the materials were selected in line with our goals.

Replicating Experiment 1, the r parameter, representing the proportion of RH use, showed strong variability between participants ($\mu^{r1} = .59$ [.51, .66], $\sigma^{r1} = .86$ [.71, 1.03], and $\mu^{r2} = .73$ [.66, .78], $\sigma^{r2} = .82$ [.69, .99], for Tests 1 and 2, respectively). Again, we observed a difference in the average r

parameters between the two choice sets, this time opposite to that in Experiment 1: Participants used the RH on average *less* often in the first than in the second test ($\Delta_r = -.14$ [–.20, –.07]). As in Experiment 1, this difference should have impacted stability only marginally.

To test the main hypothesis, we again examined the correlations between the first and second tests of the RH for both groups separately. The results showed very similar positive correlations ($\rho_{r1,r2} = .54$ [.26, .75] and $\rho_{r1,r2} = .53$ [.25, .75] for the day and week groups, respectively). This suggests that RH use is stable across time even when the choice objects differ in the two tests, ruling out the objection to Experiment 1 that stability is perhaps limited to exact replications of choices.

Notably, the between-test correlations in Experiment 2 were lower than the within-test correlations ($\rho_{r1,r2} = .90$ [.80, .96] and $\rho_{r1,r2} = .91$ [.83, .97] for the first and second tests, respectively). The former also tended to be lower than the between-test correlations observed in Experiment 1. This suggests that consistency in RH use partly depends on the similarity of (or overlap in) choice objects. To further study the influence of differences in materials, we conducted a third experiment in which we assessed the stability of RH use across *different* domains.

Experiment 3

Method

Design and procedure Participants worked on two tests of the RH (both consisting of a recognition task and a comparison task, in random order) within a single session. As in Experiment 2, we used different materials on the two tests of the RH. However, there was one important modification: The corresponding materials were drawn from two different judgment domains. The experiment again comprised two groups. In the *related group*, participants worked on two different object sets drawn from similar (although not identical) domains. In the *different group*, in contrast, the two object sets were drawn from clearly distinct domains. To maintain a high level of motivation across the lengthy experiment, participants received performance-contingent payment. In both comparison tasks, participants gained €0.03 for each correct judgment, whereas they lost €0.03 for each false judgment. However, to avoid strategy-learning effects, participants received feedback about their performance at the end of the experiment only.

Material We used different materials for the two groups. In the *related* group, participants were asked to decide on (1) the success of celebrities and (2) the success of movies, in random order. The domain of celebrities consisted of the 100 most

successful celebrities according to the Forbes List 2012 (www.forbes.com), which defined success as entertainment-related earnings plus media visibility. The domain of movies contained the 100 most successful German movies, characterized by the numbers of cinemagoers in Germany. Analogously, in the *different* group, participants were asked to decide on (1) the size of islands and (2) the success of musicians, in random order. The domain of islands included the 60 largest islands worldwide, whereas the domain of musicians involved the world's 150 most successful musicians, characterized by the numbers of records sold worldwide. To make sure that for both groups the two choice sets had similar properties (i.e., proportions of objects recognized, recognition, and knowledge validities), objects were selected on the basis of the recognition judgments of an independent prestudy. Specifically, we chose a random sample of 25 objects for each of the four domains for the recognition task. Each of these samples was then exhaustively paired, resulting in 300 trials for the comparison task.

Participants A total of 135 student participants were recruited at the University of Mannheim and randomly assigned to one of the two groups outlined above. The sample consisted of 87 women and 48 men, between 18 and 45 years of age ($M = 21.6$ years, $SD = 3.6$). All participants were native speakers or fluent in German. They received an average salary of €3.70 ($SD = 1.85$).

Results and discussion

Three participants had to be excluded from the analyses because they recognized either all but one or none of the objects. Descriptive analyses revealed that the materials were chosen in line with our goals: Participants recognized on average 13.8 celebrities ($SD = 3.3$) and 15.4 movies ($SD = 4.0$), as well as 12.6 islands ($SD = 2.6$) and 16.0 musicians ($SD = 4.0$), resulting in sufficient numbers of recognition cases for all domains. In the *related* group, the mean recognition and knowledge validities did not differ across materials ($\overline{\alpha} = .64$, $SD = .10$; $\overline{\beta} = .56$, $SD = .09$, and $\overline{\alpha} = .64$, $SD = .09$; $\overline{\beta} = .56$, $SD = .10$, for celebrities and movies, respectively) [all $ts(67) < 0.35$, $ps > .72$, $BF_{10}s < 0.14$]. In the *different* group, the mean recognition validities were similar for both domains ($\overline{\alpha} = .68$, $SD = .08$, and $\overline{\alpha} = .69$, $SD = .14$, for islands and musicians, respectively) [$t(63) = 0.64$, $p = .53$, $BF_{10} = 0.17$]. The difference in mean knowledge validities was most probably due to the choice of the materials being based on a small prestudy ($\overline{\beta} = .65$, $SD = .08$, and $\overline{\beta} = .60$, $SD = .08$, for islands and musicians, respectively) [$t(63) = 3.94$, $p < .001$, $BF_{10} = 109.0$]. Overall, the recognition and knowledge validities showed that both strategies—RH use and knowledge use—were reasonable [all $ts(67) > 4.63$, $ps < .001$, $BF_{10}s > 1,000$ for the *related* group; all $ts(63) > 10.5$, $ps < .001$, $BF_{10}s > 1,000$ for the *different* group].

Again, we found strong heterogeneity in RH use between participants in both groups ($\mu^{r1} = .77$ [.72, .82], $\sigma^{r1} = .68$ [.56, .83], and $\mu^{r2} = .83$ [.78, .86], $\sigma^{r2} = .63$ [.50, .78], for celebrities and movies, respectively; $\mu^{r1} = .69$ [.61, .76], $\sigma^{r1} = .84$ [.68, 1.04], and $\mu^{r2} = .82$ [.76, .87], $\sigma^{r2} = .76$ [.62, .95], for islands and musicians, respectively). As before, the difference in mean levels of RH use should not have influenced the results in a crucial manner ($\Delta_r = -.05$ [–.10, .001] and $\Delta_r = -.13$ [–.21, –.05], for the *related* and *different* groups, respectively).

To test the main hypothesis, the correlations between RH use in the two tasks were examined separately for each group. Overall, the results showed medium to strong correlations for both groups ($\rho_{r1,r2} = .42$ [.18, .62] and $\rho_{r1,r2} = .33$ [.08, .55], for the *related* and *different* groups, respectively), thus supporting the hypothesis of stability across domains. However, both correlations were substantially lower than the within-test correlations ($\rho_{r1,r2} = .81$ [.66, .92] for celebrities and $\rho_{r1,r2} = .77$ [.59, .90] for movies; $\rho_{r1,r2} = .89$ [.79, .96] for islands and $\rho_{r1,r2} = .81$ [.65, .92] for musicians). These differences demonstrate a potential impact of variation in domains. Also, there seems to be a downward trend in the stability coefficients, as compared to the results of Experiments 1 and 2.

In conclusion, the results of Experiments 1 to 3 support the hypothesis that RH use is relatively stable when using the same or different choice objects and also when using similar or clearly distinct domains. However, they also support the conjecture that the overall similarity of the to-be-compared decision scenarios impacts the stability of decision strategies across domains.

Another important aspect of the choice context has not been addressed so far: Experiments 1 to 3 presented choice options in verbal form only. This is a rather abstract presentation format relative to typical choice situations in everyday life. Does stability in RH use generalize to perceptually enriched, and presumably ecologically more valid, pictorial presentations of choice objects? Experiment 4 was designed to address this question.

Experiment 4

Method

Design and procedure Participants worked on two choice tasks within one session using exactly the same materials but different presentation formats. In the first task, the choice options were indicated verbally (i.e., using names), whereas in the second task they were indicated pictorially (i.e., using photos), or vice versa. Once again, participants received performance-contingent payment to maintain a high motivational level throughout the experiment, but they were informed about their overall performance only after the whole experiment was completed.

Material We used the names and pictures of the 100 most successful celebrities according to the Forbes List 2012 as choice options (cf. Exp. 3). For the recognition task, a subset of 25 objects was randomly chosen without repetition. This set was exhaustively paired, resulting in 300 pairs of objects for the comparison task. To guarantee similar properties of the materials, we selected the objects on the basis of the recognition judgments of an independent prestudy.

Participants A total of 87 student participants were recruited via posters and mailing lists at the University of Mannheim. The sample consisted of 58 women and 29 men, between 18 and 45 years of age ($M = 22.3$ years, $SD = 4.7$). All participants were native speakers or fluent in German, and they received an average salary of €3.19 ($SD = 1.36$).

Results and discussion

Descriptive analyses revealed that the materials had been chosen in line with our goals: Participants recognized on average about half of the objects ($M = 13.2$ celebrities, $SD = 3.7$, presented as names, and $M = 12.5$ celebrities, $SD = 3.7$, presented as pictures), resulting in sufficient recognition cases. Similarly, the actual mean recognition and knowledge validities showed that both the RH and knowledge use were reasonable strategies under both presentation formats ($\overline{\alpha} = .64$, $SD = .10$; $\overline{\beta} = .62$, $SD = .09$, and $\overline{\alpha} = .65$, $SD = .10$; $\overline{\beta} = .58$, $SD = .11$, for names and pictures, respectively) [$t(86) > 6.75$, $p < .001$, $BF_{10} > 1,000$].

As before, we found large individual differences in RH use, irrespective of the presentation format ($\mu^{r1} = .74$ [.68, .79], $\sigma^{r1} = .83$ [.70, 1.00], and $\mu^{r2} = .71$ [.64, .77], $\sigma^{r2} = .84$ [.70, 1.00], for names and pictures, respectively; $\Delta_r = .03$ [–.03, .09]). More importantly, RH use was stable across presentation modes, demonstrated by a strong positive correlation of $\rho_{r1,r2} = .60$ [.44, .74]. The within-test correlations were very similar and considerably higher for each format ($\rho_{r1,r2} = .91$ [.83, .96] and $\rho_{r1,r2} = .89$ [.80, .95], for names and pictures, respectively), relative to the correlations across presentation formats. Thus, stability in RH use appears to depend, at least to a certain extent, on invariance of the presentation formats.

Stability of alternative measures of RH use

To make sure that evidence on within-individual stability was not tied to a particular measure (cf. Kantner & Lindsay, 2012) or to statistical peculiarities of the *r*-model that might bias stability assessment, we replicated the main analyses using all measures of RH use previously employed in the relevant literature: the adherence rate (i.e., the proportion of cases in which the recognized object is chosen), the indices *c* (i.e., the tendency to follow

the recognition cue) and d' (i.e., the ability to discriminate cases in which recognition yields a correct vs. a false inference) derived from signal detection theory (Pachur et al., 2009), and the discrimination index (DI; Hilbig & Pohl, 2008), similar to the discriminability parameter d'. For this purpose, we calculated the respective measures for each participant and each test occasion separately and used standard methods of stability assessment (i.e., Pearson product-moment correlation coefficients), summarized in Table 1.

Briefly, we found the same pattern of results as with the r-model. In fact, the correlation coefficients for the adherence rate and the index c were comparable in size to those for the r parameter.[7] However, the correlation coefficients for the indices DI and d' are somewhat smaller. These indices have in common that they capture the ability to discriminate cases in which the RH leads to correct versus false inferences. As such, they measure the deviation from pure RH use. Uncontrolled noise factors—for instance, the overall degree of knowledge about the domain—might affect the degree of deviation from perfect RH use. Therefore, neither the lack of a linear relationship between the r parameter and these two indices (Horn et al., 2015) nor the lower within-test correlations (see Table 2 in the Appendix) and, by implication, the lower stability of DI and d' come as a surprise.

General discussion

When making decisions, people can use different strategies: These include simple strategies like the fast-and-frugal recognition heuristic (Goldstein & Gigerenzer, 2002), which assumes that decisions are based on recognition exclusively, and more costly strategies such as the integration of knowledge stored in memory, which demand more time and cognitive resources. There are two general approaches to identifying factors that influence strategy selection. On the one hand, a fertile line of research focuses on external factors—that is, situational and domain-specific variables (e.g., Bröder & Eichler, 2006; Hilbig et al., 2010; Newell & Shanks, 2004; Richter & Späth, 2006). On the other hand, a sparsely studied line of research has focused on internal factors, such as personality traits and other persistent individual characteristics (e.g., Hilbig, 2008; Pachur et al., 2009).

It has been shown repeatedly that people differ to a large extent in applying specific strategies. This heterogeneity appears to be caused by person-specific factors, independent of contextual influences (e.g., Gigerenzer & Brighton, 2009; Hilbig & Richter, 2011; Pachur et al., 2008). Among others, Shiloh,

Koren, and Zakay (2001, p. 701) observed that "individuals seem to have personal tendencies that favor the use of compensatory or non-compensatory decision strategies, which are based on personality traits." Consequently, research on personality influences is important because contextual aspects alone cannot explain individual differences in strategy selection satisfactorily. However, prior to exploring the personality determinants of RH use, temporal and cross-situational stability in RH use needs to be demonstrated as an important precondition. If people do not apply the RH in a consistent way, it will eventually turn out to be impossible to find replicable relations between RH use and individual traits.

For these reasons, we conducted four experiments to assess four different aspects of stability in RH use—namely, stability across (1) time, (2) choice objects, (3) domains, and (4) presentation formats. In all four experiments, participants worked on two tasks measuring RH use. The stability of RH use was assessed as the cross-task correlation. To account for measurement and sampling errors in individual parameter estimates of RH use, we used a hierarchical extension of the r-model applied to the two test occasions (Klauer, 2010; Matzke et al., 2015). Moreover, to ensure that the results were not limited to a particular measure of RH use, we also evaluated stability for all measures of RH use previously employed in the relevant literature (see Hilbig, 2010, for a review and comparative evaluation).

In Experiments 1 and 2, we assessed stability across time using a delay of either one day or one week between the two choice tasks. The only difference between the two experiments was that we used exactly the same choice objects in both tasks in Experiment 1 and different choice objects in Experiment 2. The results of both experiments confirmed our hypothesis that RH use is stable across time. To provide a benchmark against which to compare the correlation coefficients, we estimated the *within-test correlation*. To this end, we split the data into the first and second 150 trials and estimated the correlations in RH use between these two parts for both test occasions separately. This coefficient can be interpreted as the "baseline" stability[8] for a zero delay. In Experiment 1, the correlations of both groups were comparable to the within-test correlations in Tests 1 and 2. In Experiment 2, the correlations across test occasions were somewhat lower than the within-test correlations. However, the within-test correlations of Experiment 2 were based on exactly the same stimulus materials, whereas the between-test correlations were based on two distinct material sets. In sum, these findings suggest that stability is largely unaffected by the time

---

[7] The smaller correlations in the day group of Experiment 2 compared to the week group are due to a single participant. Excluding this participant resulted in correlation coefficients of .60 for the adherence rate and of .57 for the index c.

[8] The within-test correlations are similar in size across Experiments 2–4, where each object was repeated 24 times in the decision task. A slight decrease was found for Experiment 1, where each object was repeated only six times. This small number of repetitions might have caused a difference in choice objects between the first and the second halves of the decision task, resulting in a somewhat smaller within-test correlation than in the experiments in which the same objects were repeated 24 times.

Table 1 Pearson correlation coefficients (and 95 % confidence intervals) across the two tests of the recognition heuristic (RH) for all measures of RH use previously employed in the relevant literature

Exp.	Group	Measure of RH Use			
		Adherence Rate	c	d'	DI
1	Day group	.90 [.81, .95]	.86 [.73, .93]	.47 [.15, .70]	.39 [.05, .65]
	Week group	.72 [.49, .86]	.70 [.46, .84]	.51 [.19, .73]	.52 [.20, .74]
2	Day group	.45 [.17, .67]	.48 [.20, .69]	.12 [−.19, .41]	.42 [.13, .64]
	Week group	.55 [.30, .73]	.55 [.30, .73]	.004 [−.30, .31]	.23 [−.08, .50]
3	Related group	.49 [.28, .65]	.48 [.27, .65]	.05 [−.19, .29]	.05 [−.19, .29]
	Different group	.43 [.21, .61]	.46 [.24, .63]	.28 [.04, .49]	.20 [−.05, .42]
4	All participants	.64 [.50, .75]	.63 [.48, .74]	.32 [.12, .50]	.27 [.06, .45]

DI, discrimination index. The correlation coefficients are shown for Experiments 1–4, separately for experimental groups.

interval between measurement occasions, but might be influenced by differences in stimulus materials.

To study stability across the different materials, we systematically increased the differences in choice objects and domains in Experiments 1 to 3. Participants worked repeatedly on exactly the same choice objects in Experiment 1, on different choice objects drawn from the same domain in Experiment 2, and on objects from two distinct judgment domains that differed slightly versus substantially in Experiment 3. Medium to strong stability in RH use was found in all experiments. However, the stability coefficients tended to decrease when the differences between tasks increased: The correlations were very high and comparable to the within-test correlations when using exactly the same objects in both tests. They were lower when using different stimulus materials from the same object domain, dropped again when using objects from slightly different domains, and dropped even more when using objects from substantially different domains.

Finally, we examined stability across different presentation formats in Experiment 4, in which we presented the choice objects as names versus pictures. As before, RH use was relatively stable across tasks. However, as compared to the within-test correlations (which were similar for the two presentation formats and in line with those in the other experiments), stability was reduced slightly by a change in presentation formats.

One potential objection refers to the possibility that the observed stability in RH use was perhaps nothing but an epiphenomenon of stability in the participants' knowledge. Hilbig and Pohl (2009) have shown that more valid knowledge leads to less use of the RH. Therefore, one might hypothesize that the stability of RH use was caused by stable underlying differences in knowledge validity, with individuals high in knowledge validity using the RH less often than those with low knowledge validity. However, recall that we found stability across different domains in Experiment 3. Assuming that knowledge validity is domain-specific (i.e., uncorrelated between domains), cross-domain stability in RH use indicates that this result cannot be accounted for solely by stability in knowledge validity. Of

course, one could maintain that some aspects of knowledge are perhaps domain-general, leading to positive knowledge correlations between domains. For instance, people who have more valid knowledge concerning the domain of celebrities might also have more valid knowledge concerning movies. However, at least three aspects of our results are inconsistent with the idea that stability of RH use is caused by individual differences in knowledge (see Tables 4 to 6 in the Appendix). First, we found a reliable negative correlation between RH use and knowledge validity for one group in Experiment 2 only. Second, the retest correlations between knowledge validities either were comparable in size to the retest correlations for RH use or were even smaller and not reliable. Third, the correlations between RH use on Tests 1 and 2 were very similar in size to the partial correlations (partialing out the effect of knowledge measured on Test 1 or 2, respectively), showing that the stability of RH use is unaffected by individual differences in knowledge.

Arguing along similar lines, one might hypothesize that stability in RH use is perhaps an epiphenomenon of stability in individual recognition validities. It has repeatedly been shown that recognition validity differences between domains are positively correlated with domain-specific RH use (Gigerenzer & Goldstein, 2011; Hilbig et al., 2010; Pachur et al., 2011). However, we found no reliable positive relation between individual RH use and the individual recognition validities in any experiment (see Table 4). Also, stability in RH use is unaffected by individual differences in recognition validities, since the partial correlations closely match the zero-order test–retest correlations (see Table 6). Note that this result is not in direct conflict with the previous studies cited above, because these studies assessed recognition validity differences *between domains*, whereas we investigated recognition validity differences *between individuals*.

In sum, stability in RH use was found across time, choice objects, domains, and presentation formats to a degree similar to what has previously been found for some other trait-like

variables in judgment and decision making.[9] Moreover, the stability of RH use is not affected by individual differences in knowledge or recognition validities, suggesting that it truly reflects a specific style of decision making rather than individual differences in the information on which decisions and inferences are based.

Stability as an important precondition opens the way for exploring personality as a source of individual variation in decision-making styles. However, our results also reveal one limitation: We should not expect correlations between RH use and personality traits larger than the stability coefficients observed here. If the correlation between two tasks measuring RH use in different domains does not exceed .33 (Exp. 3, *different* group), then the correlation between a powerful personality predictor and RH use should not be expected to exceed this value, either. The insight that even powerful predictors can be expected to show moderate correlations at best provides a possible explanation for the difficulties in finding replicable relations between RH use and individual traits (Hilbig, 2008; Pachur et al., 2009).

Furthermore, our work also opens the way for exploring another rather neglected influence on decision-making styles: the *interaction* of personality and situational factors. For instance, Bröder (2003) showed that intelligence moderates adaptive use of the TTB heuristic, depending on whether or not TTB performs well in a given decision context. In our view, an analogous effect of intelligence on adaptive RH use is worth investigating. One could also think of other potential interaction effects. Hilbig (2008), for instance, suggested that participants high in neuroticism prefer RH use over knowledge use in order to avoid a diagnostic test of their abilities. If this holds, increasing the self-value relevance of the task might boost the effect of neuroticism. Furthermore, certain personality traits possibly reveal their influence only under certain situational conditions. For instance, even if impulsivity by itself is not a predictor of strategy selection (Bröder, 2012), a context condition such as time pressure might turn it into one. By contrast, strong situational influences might also eliminate the effect of personality. For instance, the lack of evidence for an association between strategy use and the need for cognition (Bröder, 2012) might originate from situational influences overshadowing personality influences. Controlling for situational influences as strictly as possible might reveal that the need for cognition is indeed an important predictor. We thus suggest using strictly neutral decision contexts (i.e., "weak situations"; cf. Mischel, 1973) if the goal is to study pure influences of personality traits on strategy use.

Moreover, following Kantner and Lindsay's (2012, 2014) analysis of individual differences in response bias, we might

ask whether RH use can be conceived of as a *cognitive trait*, meaning "an aspect of cognition that typifies an individual" (Kantner & Lindsay, 2012, p. 1164). Given the present data, we cannot answer this question now. However, we are sure that it will inspire future research. It would be interesting to analyze, for example, whether people who prefer RH use over knowledge use also favor other fast-and-frugal heuristics, such as the TTB heuristic (Gigerenzer & Goldstein, 1999). Given that the recognition and TTB heuristics share the one-reason decision-making principle, correlations between preferences for the two heuristics seem very likely. Furthermore, one could also explore whether RH use is related to other response tendencies and biases as part of "a more general, intra-individually stable decision-making heuristic" (Kantner & Lindsay, 2012, p. 1175).

In any case, one important conclusion can be drawn from the present study: The likelihood of RH use is not only influenced by situational determinants that affect the costs and benefits of RH use (Erdfelder, Küpper-Tetzel, & Mattern, 2011; Hilbig et al., 2010; McCloy, Beaman, Frosch, & Goddard, 2010; Oppenheimer, 2003; Pachur & Biele, 2007; Pohl, Erdfelder, Hilbig, Liebke, & Stahlberg, 2013; Schooler & Hertwig, 2005). As we have shown in the present research, it is also influenced by relatively stable individual tendencies favoring either RH use or the integration of further knowledge. Thus, our work contributes to a new line of research on the cognitive and personality traits underlying RH use, aiming at a comprehensive theory that integrates situational and personality determinants of decision strategies.

Appendix

Table 2 Main results concerning RH use in Experiment 1, using the original recognition judgments of each session

	Day Group		Week Group	
	Test 1	Test 2	Test 1	Test 2
Mean	.76 [.71, .80]	.73 [.66, .79]	.78 [.73, .83]	.69 [.62, .75]
SD	.39 [.27, .55]	.48 [.34, .67]	.39 [.27, .56]	.43 [.30, .60]
Correlation	.67 [.34, .88]		.58 [.23, .84]	

Means, standard deviations, and correlation coefficients (with 95 % Bayesian credible intervals) are measured via the hierarchical *r*-model and shown separately for experimental groups and the two tests of the RH. *Test 1* refers to the initial test of the RH, whereas *Test 2* refers to the second test of the RH, done one day or one week later.

[9] To illustrate, a stable response bias was shown using different time intervals ($\rho \in$ [.67, .73]) and presentation formats ($\rho \in$ [.33, .81]; Kantner & Lindsay, 2012). Similarly, delay discounting was found to be stable across time ($\rho \in$ [.71, .91]), stimulus materials ($\rho \in$ [.18, .90]), and presentation formats ($\rho \in$ [.44, .83]; Odum, 2011).

Table 3 Within-test correlation coefficients (and 95 % confidence intervals) for all measures of RH use previously employed in the relevant literature

Exp.	Material	Measure of RH Use			
		Adherence Rate	c	d'	DI
1	Test 1	.80 [.69, .87]	.79 [.67, .86]	.07 [–.18, .31]	.33 [.09, .53]
	Test 2	.84 [.75, .90]	.76 [.63, .84]	.40 [.18, .59]	.40 [.18, .59]
2	Test 1	.91 [.86, .94]	.91 [.86, .94]	.50 [.32, .64]	.63 [.47, .74]
	Test 2	.90 [.85, .93]	.91 [.87, .94]	.65 [.50, .76]	.77 [.66, .84]
3	Celebrities	.88 [.81, .92]	.88 [.82, .93]	.20 [–.04, .42]	.21 [–.03, .43]
	Movies	.82 [.72, .88]	.80 [.69, .87]	.38 [.15, .57]	.31 [.08, .51]
	Islands	.93 [.89, .96]	.92 [.87, .95]	.38 [.15, .57]	.49 [.28, .66]
	Musicians	.86 [.78, .91]	.87 [.79, .92]	.47 [.25, .64]	.58 [.39, .73]
4	Names	.92 [.89, .95]	.93 [.89, .95]	.46 [.27, .61]	.44 [.26, .60]
	Pictures	.91 [.87, .94]	.91 [.86, .94]	.55 [.38, .68]	.55 [.38, .68]

DI, discrimination index. The correlation coefficients are shown for Experiments 1–4, separately for the two tests of the RH. For Experiments 1 and 2, *Test 1* refers to the initial test of the RH, whereas *Test 2* refers to the second test of the RH, performed one day or one week later. Within-test correlations are estimated using the Spearman–Brown-corrected Pearson correlation coefficient.

Table 4 Correlation coefficients (with 95 % Bayesian credible intervals) for the correlations between RH use and recognition validity and between RH use and knowledge validity, separately for the two tests of the RH

Exp.	Group	Correlation Between RH Use and			
		Recognition Validity		Knowledge Validity	
		Test 1	Test 2	Test 1	Test 2
1	Day group	.35 [–.09, .71]	.37 [–.05, .71]	.13 [–.05, .55]	.29 [–.14, .68]
	Week group	.41 [–.01, .73]	.31 [–.12, .67]	.25 [–.48, .80]	.38 [–.13, .79]
2	Day group	–.06 [–.40, .29]	–.16 [–.48, .19]	**–.35 [–.62, –.02]**	**–.36 [–.63, –.04]**
	Week group	–.01 [–.10, .35]	.06 [–.32, .42]	–.03 [–.16, .31]	.01 [–.33, .35]
3	Related group	–.02 [–.28, .25]	.16 [–.11, .42]	.36 [.09, .60]	.07 [–.22, .36]
	Different group	.23 [–.05, .48]	.22 [–.03, .46]	.19 [–.11, .48]	.19 [–.12, .47]
4	All participants	.01 [–.22, .25]	–.19 [–.40, .40]	.58 [.36, .76]	.22 [–.01, .45]

The correlation coefficients are shown for Experiments 1–4, separately for experimental groups and the two tests of the RH. Boldface indicates significant correlations, shown by credible intervals that do not include 0. For Experiments 1 and 2, *Test 1* refers to the initial test of the RH, whereas *Test 2* refers to the second test of the RH, performed one day or one week later. For Experiment 3, *Tests 1 and 2* refer to the two different domains that were used as materials (celebrities vs. movies and islands vs. musicians). For Experiment 4, *Tests 1 and 2* refer to the two presentation formats (names vs. pictures). Recognition and knowledge validities are assessed via the *a* and *b* parameters of the *r*-model, respectively. Correlations are estimated using the hierarchical *r*-model.

Table 5 Correlation coefficients (with 95 % Bayesian credible intervals) for recognition and knowledge validities across tests

Exp.	Group	Test-Retest Correlations	
		Recognition Validity	Knowledge Validity
1	Day group	.81 [.57, .95]	.67 [.29, .91]
	Week group	.88 [.71, .97]	.48 [−.24, .88]
2	Day group	−.10 [−.47, .27]	.53 [.21, .78]
	Week group	.05 [−.08, .43]	.51 [.41, .78]
3	Related group	.38 [.11, .62]	.06 [−.28, .38]
	Different group	.43 [.18, .65]	.28 [−.10, .61]
4	All participants	.16 [−.08, .38]	.70 [.46, .85]

The correlation coefficients are shown for Experiments 1–4, separately for experimental groups. Recognition and knowledge validities are assessed via the *a* and *b* parameters of the *r*-model, respectively. Correlations are estimated using the hierarchical *r*-model.

Table 6 Comparison between zero-order and partial correlation coefficients (with 95 % Bayesian credible intervals) for RH use across tests, partialing out the effects of the recognition validities of Tests 1 and 2 and the knowledge validities of Tests 1 and 2, respectively

Exp.	Group	Zero-Order Correlation	Partial Correlation Controlling for			
			Recognition Validity		Knowledge Validity	
			Test 1	Test 2	Test 1	Test 2
1	Day group	.80 [.55, .94]	.78 [.52, .94]	.78 [.54, .94]	.81 [.57, .95]	.79 [.55, .94]
	Week group	.71 [.39, .91]	.66 [.33, .89]	.66 [.32, .89]	.66 [.24, .90]	.66 [.30, .90]
2	Day group	.54 [.26, .75]	.55 [.28, .75]	.54 [.26, .75]	.48 [.19, .71]	.46 [.16, .69]
	Week group	.53 [.25, .75]	.53 [.24, .76]	.55 [.28, .77]	.53 [.25, .75]	.54 [.25, .76]
3	Related group	.42 [.17, .63]	.42 [.18, .62]	.40 [.15, .61]	.36 [.10, .59]	.42 [.17, .63]
	Different group	.33 [.08, .55]	.33 [.08, .55]	.34 [.09, .56]	.28 [.003, .52]	.29 [.03, .52]
4	All participants	.60 [.43, .74]	.60 [.44, .74]	.59 [.43, .73]	.48 [.26, .67]	.57 [.39, .72]

The zero-order and partial correlation coefficients are shown for Experiments 1–4, separately for experimental groups. For Experiments 1 and 2, *Test 1* refers to the initial test of the RH, whereas *Test 2* refers to the second test of the RH, performed one day or one week later. For Experiment 3, *Tests 1 and 2* refer to the two different domains that were used as materials (celebrities vs. movies and islands vs. musicians). For Experiment 4, *Tests 1 and 2* refer to the two presentation formats (names vs. pictures). Recognition and knowledge validities are assessed via the *a* and *b* parameters of the *r*-model, respectively. Correlations and partial correlations are estimated using the hierarchical *r*-model.

References

Aminoff, E. A., Clewett, D., Freeman, S., Frithsen, A., Tipper, C., Johnson, A., . . . Miller, M. B. (2012). Individual differences in shifting decision criterion: A recognition memory study. *Memory & Cognition, 40,* 1016–1030. doi:10.3758/s13421-012-0204-6

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57–86. doi:10.3758/BF03210812

Bröder, A. (2003). Decision making with the "Adaptive Toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 611–625. doi:10.1037/0278-7393.29.4.611

Bröder, A. (2012). The quest for take the best—Insights and outlooks from experimental research. In P. M. Todd, G. Gigerenzer, & the ABC Research Group (Eds.), *Ecological rationality: Intelligence in the world* (pp. 216–240). New York, NY: Oxford University Press.

Bröder, A., & Eichler, A. (2006). The use of recognition information and additional cues in inferences from memory. *Acta Psychologica, 121,* 275–284. doi:10.1016/j.actpsy.2005.07.001

Bröder, A., & Newell, B. R. (2008). Challenging some common beliefs: Empirical work within the adaptive toolbox metaphor. *Judgment and Decision Making, 3,* 205–214. **Retrieved from http://journal.sjdm.org/bn2.pdf**

Castela, M., Kellen, D., Erdfelder, E., & Hilbig, B. E. (2014). The impact of subjective recognition experiences on recognition heuristic use: A multinomial processing tree approach. *Psychonomic Bulletin & Review, 21,* 1131–1138. doi:10.3758/s13423-014-0587-4

Couch, A., & Keniston, K. (1960). Yeasayers and naysayers: Agreeing response set as a personality variable. *Journal of Abnormal and Social Psychology, 60,* 151–174. doi:10.1037/h0040372

Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models. *Journal of Psychology, 217,* 108–124. doi:10.1027/0044-3409.217.3.108

Erdfelder, E., Küpper-Tetzel, C. E., & Mattern, S. D. (2011). Threshold models of recognition and the recognition heuristic. *Judgment and Decision Making, 6,* 7–22. **Retrieved from http://journal.sjdm.org/11/rh13/rh13.pdf**

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. Boca Raton, FL: Chapman & Hall.

Gigerenzer, G., & Brighton, H. (2009). *Homo heuristicus*: Why biased minds make better inferences. *Topics in Cognitive Science, 1,* 107–143. doi:10.1111/j.1756-8765.2008.01006.x

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The Take the Best Heuristic. In G. Gigerenzer, P. M. Todd, & The ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 75–95). New York, NY: Oxford University Press.

Gigerenzer, G., & Goldstein, D. G. (2011). The recognition heuristic: A decade of research. *Judgment and Decision Making, 6,* 100–121. **Retrieved from http://journal.sjdm.org/11/rh15/rh15.pdf**

Glöckner, A., & Bröder, A. (2011). Processing of recognition information and additional cues: A model-based analysis of choice, confidence, and response time. *Judgment and Decision Making, 6,* 23–42. **Retrieved from http://journal.sjdm.org/11/rh4/rh4.pdf**

Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition, 123,* 21–32. doi:10.1016/j.cognition.2011.12.002

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109,* 75–90. doi:10.1037/0033-295X.109.1.75

Hilbig, B. E. (2008). Individual differences in fast-and-frugal decision making: Neuroticism and the recognition heuristic. *Journal of Research in Personality, 42,* 1641–1645. doi:10.1016/j.jrp.2008.07.001

Hilbig, B. E. (2010). Precise models deserve precise measures: A methodological dissection. *Judgment and Decision Making, 5,* 272–284. **Retrieved from http://journal.sjdm.org/10/rh5/rh5.pdf**

Hilbig, B. E., Erdfelder, E., & Pohl, R. F. (2010). One-reason decision making unveiled: A measurement model of the recognition heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 123–134. doi:10.1037/a0017518

Hilbig, B. E., Erdfelder, E., & Pohl, R. F. (2012). A matter of time: Antecedents of one-reason decision making based on recognition. *Acta Psychologica, 141,* 9–16. doi:10.1016/j.actpsy.2012.05.006

Hilbig, B. E., Michalkiewicz, M., Castela, M., Pohl, R. F., & Erdfelder, E. (2015). Whatever the cost? Information integration in memory-based inferences depends on cognitive effort. *Memory & Cognition, 43,* 659–671. doi:10.3758/s13421-014-0493-z

Hilbig, B. E., & Pohl, R. F. (2008). Recognizing users of the recognition heuristic. *Experimental Psychology, 55,* 394–401. doi:10.1027/1618-3169.55.6.394

Hilbig, B. E., & Pohl, R. F. (2009). Ignorance- versus evidence-based decision making: A decision time analysis of the recognition heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1296–1305. doi:10.1037/a0016565

Hilbig, B. E., Pohl, R. F., & Bröder, A. (2009). Criterion knowledge: A moderator of using the recognition heuristic? *Journal of Behavioral Decision Making, 22,* 510–522. doi:10.1002/bdm.644

Hilbig, B. E., & Richter, T. (2011). *Homo heuristicus* outnumbered: Comment on Gigerenzer and Brighton (2009). *Topics in Cognitive Science, 3,* 187–196. doi:10.1111/j.1756-8765.2010.01123.x

Hilbig, B. E., Scholl, S. G., & Pohl, R. F. (2010). Think or blink—Is the recognition heuristic an "intuitive" strategy? *Judgment and Decision Making, 5,* 300–309. **Retrieved from http://journal.sjdm.org/10/rh6/rh6.pdf**

Horn, S. S., Pachur, T., & Mata, R. (2015). How does aging affect recognition-based inference? A hierarchical Bayesian modeling approach. *Acta Psychologica, 154,* 77–85. doi:10.1016/j.actpsy.2014.11.001

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Kantner, J., & Lindsay, D. S. (2012). Response bias in recognition memory as a cognitive trait. *Memory & Cognition, 40,* 1163–1177. doi:10.3758/s13421-012-0226-0

Kantner, J., & Lindsay, D. S. (2014). Cross-situational consistency in recognition memory response bias. *Psychonomic Bulletin & Review, 21,* 1272–1280. doi:10.3758/s13423-014-0608-3

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika, 75,* 70–98. doi:10.1007/S11336-009-9141-0

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. New York, NY: Cambridge University Press.

Lunn, D., Spiegelhalter, D., Thomas, A., & Best, N. (2009). The BUGS project: Evolution, critique, and future directions. *Statistics in Medicine, 28,* 3049–3067. doi:10.1002/sim.3680

Marewski, J. N., Gaissmaier, W., Schooler, L. J., Goldstein, D. G., & Gigerenzer, G. (2010). From recognition to decisions: Extending and testing recognition-based models for multialternative inference. *Psychonomic Bulletin & Review, 17,* 287–309. doi:10.3758/PBR.17.3.287

Matzke, D., Dolan, C. V., Batchelder, W. H., & Wagenmakers, E.-J. (2015). Bayesian estimation of multinomial processing tree models with heterogeneity in participants and items. *Psychometrika, 80,* 205–235. doi:10.1007/s11336-013-9374-9

Matzke, D., Lee, M. D., & Wagenmakers, E.-J. (2013). Multinomial processing trees. In M. D. Lee & E.-J. Wagenmakers (Eds.), *Bayesian cognitive modeling: A practical course* (pp. 187–195). New York, NY: Cambridge University Press.

McCloy, R. A., Beaman, C. P., Frosch, C. A., & Goddard, K. (2010). Fast and frugal framing effects? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1043–1052. doi:10.1037/a0019693

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Personality Psychological Review, 80,* 252–283. doi:10.1037/h0035002

Moshagen, M. (2010). multiTree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods, 42,* 42–54. doi:10.3758/BRM.42.1.42

Newell, B. R., & Fernandez, D. (2006). On the binary quality of recognition and the inconsequentiality of further knowledge: Two critical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19,* 333–346. doi:10.1002/bdm.531

Newell, B. R., & Shanks, D. R. (2004). On the role of recognition in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 923–935. doi:10.1037/0278-7393.30.4.923

Odum, A. L. (2011). Delay discounting: Trait variable? *Behavioral Processes, 87,* 1–9. doi:10.1016/j.beproc.2011.02.007

Oppenheimer, D. M. (2003). Not so fast! (and not so frugal!): Rethinking the recognition heuristic. *Cognition, 90,* B1–B9. doi:10.1016/S0010-0277(03)00141-0

Pachur, T., & Biele, G. (2007). Forecasting from ignorance: The use and usefulness of recognition in lay predictions of sports events. *Acta Psychologica, 125,* 99–116. doi:10.1016/j.actpsy.2006.07.002

Pachur, T., Bröder, A., & Marewski, J. N. (2008). The recognition heuristic in memory-based inference: Is recognition a non-compensatory cue? *Journal of Behavioral Decision Making, 21,* 183–210. doi:10.1002/bdm.581

Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 983–1002. doi:10.1037/0278-7393.32.5.983

Pachur, T., Mata, R., & Schooler, L. J. (2009). Cognitive aging and the adaptive use of recognition in decision making. *Psychology and Aging, 24,* 901–915. doi:10.1037/a0017211

Pachur, T., Todd, P. M., Gigerenzer, G., Schooler, L. J., & Goldstein, D. G. (2011). The recognition heuristic: A review of theory and tests. *Frontiers in Psychology, 2*(147), 1–14. doi:10.3389/fpsyg.2011.00147

Pohl, R. F. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19,* 251–271. doi:10.1002/bdm.522

Pohl, R. F. (2011). On the use of recognition in inferential decision making: An overview of the debate. *Judgment and Decision Making, 6,* 423–438. **Retrieved from** http://journal.sjdm.org/11/rh19/rh19.pdf

Pohl, R. F., Erdfelder, E., Hilbig, B. E., Liebke, L., & Stahlberg, D. (2013). Effort reduction after self-control depletion: The role of cognitive resources in use of simple heuristics. *Journal of Cognitive Psychology, 25,* 267–276. doi:10.1080/20445911.2012.758101

Pohl, R. F., von Massow, F., & Beckmann, B. (2015). *Developmental differences in using an ecologically valid decision strategy: The case of the recognition heuristic.* Manuscript submitted for publication.

Richter, T., & Späth, P. (2006). Recognition is used as one cue among others in judgment and decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 150–162. doi:10.1037/0278-7393.32.1.150

Roberts, B. W. (2009). Back to the future: Personality and assessment and personality development. *Journal of Research in Personality, 43,* 137–145. doi:10.1016/j.jrp.2008.12.015

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16,* 225–237. doi:10.3758/PBR.16.2.225

Scheibehenne, B., & Bröder, A. (2007). Predicting Wimbledon 2005 tennis results by mere player name recognition. *International Journal of Forecasting, 23,* 415–426. doi:10.1016/j.ijforecast.2007.05.006

Scheibehenne, B., & Pachur, T. (2015). Using Bayesian hierarchical parameter estimation to assess the generalizability of cognitive models of choice. *Psychonomic Bulletin & Review, 22,* 391–407. doi:10.3758/s13423-014-0684-4

Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review, 112,* 610–628. doi:10.1037/0033-295X.112.3.610

Shiloh, S., Koren, S., & Zakay, D. (2001). Individual differences in compensatory decision-making style and need for closure as correlates of subjective decision complexity and difficulty. *Personality and Individual Differences, 30,* 699–710. doi:10.1016/S0191-8869(00)00073-8

Sturtz, S., Ligges, U., & Gelman, A. (2005). R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software, 12,* 1–16. **Retrieved from** www.jstatsoft.org/v12/i03/paper

Witkin, H. A., Goodenough, D. R., & Karp, S. A. (1967). Stability of cognitive style from childhood to young adulthood. *Journal of Personality and Social Psychology, 7,* 291–300. doi:10.1037/h0025070

Yechiam, E., & Busemeyer, J. R. (2008). Evaluating generalizability and parameter consistency in learning models. *Games and Economic Behaviour, 63,* 370–394. doi:10.1016/j.geb.2007.08.011

Do smarter people make better decisions? The influence of intelligence on adaptive use
of the recognition heuristic

Martha Michalkiewicz, Katja Arden, and Edgar Erdfelder

University of Mannheim

Abstract

Within the adaptive toolbox approach, it has repeatedly been shown that, on average, people tend to adapt their decision strategies to the decision context. However, it remains unclear whether individuals systematically differ in their ability to successfully adapt to the situation. We addressed this question with respect to the fast-and-frugal recognition heuristic (RH). When deciding between recognized and unrecognized choice objects, individuals can base their decisions solely on recognition, as predicted by the RH, or they can integrate further knowledge retrieved from memory. Since intelligence has been conceived as the ability to successfully adapt to different situations, we expected intelligence to influence the degree of adaptive use of the RH. To test this hypothesis, we first reanalyzed a study that assessed individual RH-use in a decision domain for which RH-use is known to be very efficient. As expected, RH-use increased with general intelligence. Next, we designed an experiment addressing individual RH-use in two new decision domains, one domain for which RH-use was less efficient than knowledge integration and another domain for which both strategies were about equally efficient. Moreover, we tested whether fluid or crystallized intelligence best predicts adaptive use of the RH. RH-use was found to decrease with fluid but not crystallized intelligence when RH-use was less efficient than knowledge integration. In contrast, there was no association between either type of intelligence and RH-use when none of the two strategies was optimal. Hence, adaptive use versus non-use of the RH appears to be moderated by fluid intelligence.

Keywords: Adaptive decision making, recognition heuristic, intelligence, multinomial processing tree models, hierarchical Bayesian modeling

Do smarter people make better decisions? The influence of intelligence on adaptive use of the recognition heuristic

Introduction

Individuals display a great deal of adaptivity in decision making. As such, individuals select decision strategies in accordance with both their own current processing resources and the characteristics of the decision context (e.g., Gigerenzer & Selten, 2001; Payne, Bettman, & Johnson, 1988, 1993; Simon, 1956, 1990). Building upon this insight, Gigerenzer, Todd, and the ABC Research Group (1999) suggested the *adaptive toolbox*. The basic idea is that people possess a repertoire of decision strategies, termed heuristics, to solve the decision problems they face. Heuristics address the limited cognitive resources in a specific context by requiring only a minimum of knowledge and information processing. Moreover, heuristics are domain-specific, that is, tailored to a specific type of decision problem, and ecologically rational to the degree they match the structure of the environment (Gigerenzer et al., 1999).

In recent years, several fast-and-frugal heuristics have been proposed as part of the adaptive toolbox. One of the most prominent examples is the recognition heuristic (RH; Goldstein & Gigerenzer, 1999, 2002). According to the RH, decision makers should base their choice solely on recognition while ignoring any additional information. Alternatively, they can integrate knowledge retrieved from memory. For instance, when asked to decide which of two cities is more populous (with one city recognized and the other not) a decision maker can either simply chose the city he or she recognizes, or use knowledge retrieved from memory. To illustrate, the decision maker might know that the recognized city has an international airport and that cities with an international airport are most often more populous than cities without one.

Previous research has shown that people generally tend to apply the RH adaptively. For example, decision makers adapt to the validity of the recognition cue, that is, to the correlation between recognition and the criterion of interest (e.g., the population of cities). On average, they rely more on the RH when recognition validity is high compared to when it is low (e.g., Castela, Kellen, Erdfelder, & Hilbig, 2014; Hilbig, Erdfelder, & Pohl, 2010; Pachur, Mata, & Schooler, 2009; Pohl, 2006;

Scheibehenne & Bröder, 2007). In contrast, decision makers apply the RH less when valid knowledge is easily available and can be integrated without much effort (e.g., Bröder & Eichler, 2006; Glöckner & Bröder, 2011; Hilbig, Michalkiewicz, Castela, Pohl, & Erdfelder, 2015; Hilbig, Pohl, & Bröder, 2009; Newell & Fernandez, 2006; Richter & Späth, 2006). Furthermore, decision makers use the RH in accordance with current constraints evoked by the context, such as time pressure (e.g., Hilbig, Erdfelder, & Pohl, 2012; Pachur & Hertwig, 2006).

Despite this massive evidence, one important question remained largely unanswered so far: Do people differ in their ability to adapt to a given decision context? And if so, which cognitive factors might explain these differences? In other words, does the decision context interact with individual characteristics? This question is of special importance as individuals' decisions are not only largely affected by situational factors, but also by individual traits (e.g., Hilbig, 2008; Pohl, von Massow, & Beckmann, 2016) and cognitive resources (e.g., Pohl, Erdfelder, Hilbig, Liebke, & Stahlberg, 2013). As already emphasized by Gigerenzer and Gaissmaier (2011) and Pachur, Bröder, and Marewski (2008), only the inclusion of both situational and individual factors into the theory of decision making will result in a comprehensive theory and a profound understanding thereof. Recently, Michalkiewicz and Erdfelder (2016) opened the way for investigating the joint effect of the decision context and cognitive abilities on adaptive use of the RH by showing both temporal and cross-situational stability in RH-use.

Building upon these results, the current paper addresses the effect of what can be considered the most fundamental cognitive resource, namely, intelligence. Neisser et al. (1996, p. 77) characterized human intelligence as the "ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought". Sternberg and Salter (1982, p. 3) simply defined intelligence as "goal-directed adaptive behavior". Obviously, the common understanding of human intelligence focuses on adaptation as a core feature. In fact, there are several ways in which intelligence could influence strategy use, for example, by affecting (1) the available strategy repertoire, (2) strategy selection, or (3) strategy execution (Lemaire, 2010). The present paper focuses on fluid and crystallized intelligence, the two major second-factors of human intelligence (Cattell, 1963), and their influence on strategy selection.

Fluid intelligence has been conceived as the capacity to think logically and solve problems in novel situations, independent of acquired knowledge (Horn & Cattell, 1966, 1967). Among others, fluid intelligence includes such abilities as pattern recognition, abstract reasoning, and problem solving. Thus, fluid intelligence can be expected to facilitate identifying the statistical structure of the environment (hence, the validity of the recognition cue compared to the validity of knowledge) and to identify the appropriate decision strategy accordingly (Horn, 1991; Mata, Schooler, & Rieskamp, 2007). By implication, fluid intelligence should foster adaptive use versus non-use of the RH. In contrast, crystallized intelligence has been conceived as the ability to use skills, knowledge, and experience acquired as a product of educational and cultural experience (Horn & Cattell, 1966, 1967). Crystallized intelligence might also foster adaptive RH-use because crystallized intelligence is associated with better knowledge about the fit between environments and potential strategies (Mata et al., 2007) and also with more extensive experience concerning strategy selection (Horn & Cattell, 1966). However, this benefit of crystallized intelligence is unlikely to emerge in any situation; it seems more plausible that it is limited to contexts where the environmental structure is well-known or can easily be derived from past experience.

Past research has revealed some support for the notion that fluid and crystallized intelligence affect adaptive use of decision strategies. For example, Bröder (2003) discovered that adaptive use of the take-the-best heuristic (TTB; Gigerenzer & Goldstein, 1999) was moderated by intelligence. In one study, participants scoring higher on fluid intelligence were more often classified as users of the more adequate strategy as determined by the payoff structure of the environment. In a second study, an influence of crystallized intelligence was also apparent, although not statistically significant. Furthermore, Mata et al. (2007) investigated the effect of age differences on adaptive use of TTB. They observed that young adults outperformed older adults in adaptive use of the TTB heuristic and explained this difference by an age-related cognitive decline (i.e., a decrease in fluid intelligence). Similarly, Pachur et al. (2009) showed that both young and older adults adjusted RH-use between environments with high and low recognition validities. However, older adults failed to abandon the RH when necessary, that is, they tended to use the RH even in contexts for which this heuristic was inappropriate. Again, these age-related constraints were explained by fluid intelligence (for a reanalysis see Horn, Pachur, & Mata, 2015).

Based on these insights, we hypothesize that intelligence will lead to better adaptation to a given situation and thus to more use of the smarter, that is, more valid, decision strategy (RH vs. knowledge integration) in the respective decision context. To test this hypothesis, we first reanalyzed a study by Hilbig (2008). In particular, we assessed whether general intelligence will be associated with more RH-use when recognition validity[1] is substantially larger than knowledge validity[2]. If true, this result will represent a first indication that intelligence moderates adaptive RH-use. However, this finding alone would not prove that intelligence affects adaptive use of the RH. A possible alternative explanation would be that RH-use perhaps generally increases with intelligence, irrespective of the decision context. Moreover, it is not clear which component or components of intelligence affect adaptive RH-use. To test between these two alternative explanations and to identify which component of general intelligence is responsible for successful adaptation to the situation, we additionally assessed the effect of fluid and crystallized intelligence on adaptive RH-use in two novel decision scenarios. First, we analyzed whether fluid and crystallized intelligence will be associated with less RH-use when recognition validity is substantially smaller than knowledge validity (*knowledge condition*) – opposite to the scenario studied by Hilbig (2008). Second, we analyzed whether both fluid and crystallized intelligence will have little effect on strategy selection when recognition and knowledge validities do not differ (*neutral condition*). Studying these scenarios will provide evidence for or against an effect of intelligence on adaptive RH-use. Furthermore, it will offer insights whether fluid or crystallized intelligence is more important for adaptive strategy selection.

Methodological preliminaries

Task

The standard paradigm to investigate the RH consists of a recognition task and a paired-comparison task (e.g., Goldstein & Gigerenzer, 2002; Hilbig & Pohl, 2009; Pachur et al., 2009). In the recognition task, participants are asked to provide yes-no

---

[1] Recognition validity is defined as the proportion of recognition pairs (i.e., one object recognized, the other not recognized) where choice of the recognized object represents the correct choice (Goldstein & Gigerenzer, 2002, p. 78).

[2] Knowledge validity is defined as the proportion of correct choices based on all pairs where both objects are recognized (Goldstein & Gigerenzer, 2002, p. 78).

recognition judgments for a set of objects drawn from a decision domain like the worlds' most populous cities. For the paired-comparison task, this set of objects is exhaustively paired. For each pair of object, participants have to decide which of the two objects has the higher value with respect to the criterion of interest (e.g., which city is more populous).

Model

We analyzed the data with the most widely used method for assessing RH-use, the r-model (see Figure 1; Hilbig et al., 2010), a multinomial processing tree (MPT) model (Batchelder & Riefer, 1999; Erdfelder et al., 2009), whose name-giving *r* parameter can be directly interpreted as the proportion of pure RH-use. The complimentary probability *1 - r*, by implication, can be interpreted as the proportion of integrating further knowledge. Compared to other measures of the RH this model offers the advantage that RH-use is assessed free of confounds with knowledge integration. For a comprehensive overview of the advantages of the r-model, see Hilbig (2010).

Specifically, we extended a hierarchical implementation of the r-model (Michalkiewicz & Erdfelder, 2016), based on the latent-trait approach to MPT models (Klauer, 2010), by including intelligence as a predictor of RH-use into the model. This hierarchical model has two main advantages: First, it assumes that individual parameters stem from group-level distributions, estimated along with individual-level parameters. Using this hierarchical framework, individual parameters are estimated more reliably compared to non-hierarchical analyses by making use of the underlying group-level structure (Gelman, Carlin, Stern, & Rubin, 2004). Second, the influence of intelligence on RH-use is assessed by directly incorporating intelligence measures into the estimation of RH-use in terms of a regression.[3] The estimated regression coefficients are thus adjusted for the uncertainty of the individual estimates of RH-use. For an introduction to hierarchical MPT models see Matzke, Dolan, Batchelder, and Wagenmakers (2015). An implementation of the latent-trait approach in other MPT models can be found in Arnold, Bayen, and Smith (2015).

---

[3] In contrast, conventional analyses consist of two-step procedures of estimating model parameters separately for each participant first, followed by regression analyses of RH-use on personality measures' scores (cf. Arnold, Bayen, & Böhm, 2014; Hilbig, 2008).

Figure 2 illustrates the hierarchical latent-trait r-model including general intelligence as a predictor of RH-use. In particular, individual RH-use r_i ($i = 1, …, I$) is modeled in a probit transformed parameter space as a linear combination of the group mean μ^r, a multiplicative scale parameter ξ^r, and individual deviations from the group mean δ_i^r. To test our hypotheses, we additionally incorporate general intelligence test scores, Int_i, with a regression coefficient β^{Int} into the estimation of individual RH-use: $\phi^{-1}(r_i) = \mu^r + \xi^r \cdot \delta_i^r + \beta^{Int} \cdot Int_i$. Generalizing the model to include two intelligence measures, $Fluid_i$ and $Cryst_i$, together with their regression coefficients $\hat{\beta}^{Fluid}$ and $\hat{\beta}^{Cryst}$, is straightforward.

To ensure that the regression coefficients are comparable, we standardized all intelligence test scores. We performed all analyses within the Bayesian framework using OpenBUGS (Lunn, Spiegelhalter, Thomas, & Best, 2009) through R2WinBUGS (Sturtz, Ligges, & Gelman, 2005). Specifically, we ran three chains with 500,000 iterations each using a thinning rate of 10 and discarded the first 100,000 iterations as burn-in period. We verified satisfactory convergence of the three chains ($\hat{R} < 1.01$; Gelman et al., 2004) and sufficiently large effective sample sizes (Kruschke, 2014). Core results are the means of the posterior distributions along with their 95% Bayesian credible intervals (BCI) for group-level recognition validity $\hat{\mu}^a$, knowledge validity $\hat{\mu}^b$, RH-use $\hat{\mu}^r$, and the variation in RH-use across subjects $\hat{\sigma}^r$. Thereby, the BCIs represent the precision of these parameter estimates. Most importantly, we are interested in the effect of intelligence on RH-use, in particular, of general intelligence $\hat{\beta}^{Int}$ in the reanalysis of Hilbig's (2008) data, and of fluid and crystallized intelligence, $\hat{\beta}^{Fluid}$ and $\hat{\beta}^{Cryst}$, in the new experiment reported here. In this case, the BCIs indicate whether the effect can be considered meaningful, that is, significantly different from zero.

Reanalysis of Hilbig's (2008) data

The original purpose of Hilbig's (2008) study was to investigate the effect of neuroticism on RH-use, while assessing general intelligence as a potential confounding variable. Participants first completed a battery of personality tests, including the Berliner Intelligenz-Struktur-Test (BIS; Jäger, Süß, & Beauducel, 1997) as a measure of general intelligence, and then worked on the standard paradigm of RH-use described above. To test for an influence of general intelligence on RH-use, we applied the hierarchical latent-trait r-model described above to the data of Hilbig (2008). Our

reanalysis shows that the decision domain used by Hilbig (i.e., city-size comparisons for the world's most populous cities) favors RH-use over knowledge-use: The mean recognition validity is clearly higher than the mean knowledge validity ($\hat{\mu}^a$ = .74 [.70; .77], $\hat{\mu}^b$ = .55 [.51; .59]; $\hat{\Delta}_{a\text{-}b}$ = .19 [.13; .25]). In line with this finding, decision makers preferred RH-use over knowledge integration ($\hat{\mu}^r$ = .73 [.64; .81]) on average. Also, the average proportion of RH-use was comparable in size to previous studies using the same material (e.g., Hilbig et al., 2010). Most importantly, there was a positive effect of general intelligence on RH-use ($\hat{\beta}^{Int}$ = 0.31 [0.05, 0.59]). Thus, participants higher in intelligence prefer the smarter, that is, more valid decision strategy when using the RH in the city-size comparison context.

Experiment

Methods

Design and Materials. We assessed use of the RH using the standard paradigm described before. Specifically, to facilitate finding an influence of crystallized intelligence we used easily accessible and common materials: The 100 most successful celebrities according to the Forbes List 2012 (www.forbes.com), which defines success as entertainment-related earnings plus media visibility. To test our core hypothesis, we manipulated the difference between recognition and knowledge validities in the paired-comparison task between participants in two conditions.[4] In particular, we asked for celebrities' age in the knowledge condition and for celebrities' success in the neutral condition. The results of an independent prestudy had previously shown that, in the knowledge condition, the mean recognition validity is significantly smaller than the mean knowledge validity ($\overline{\alpha}$ = .56, SD = .07, $\overline{\beta}$ = .77, SD = .12; $t(29)$ = 7.99, $p < .001$, $BF_{10} > 1000$)[5], whereas both validities are very similar in the neutral condition ($\overline{\alpha}$ = .64, SD = .08, $\overline{\beta}$ = .61, SD = .12; $t(29)$ = 1.18, p = .25, BF_{10} = 0.36).

We used standard measures to assess fluid and crystallized intelligence (Horn, 1991): A short form of the Advanced Progressive Matrices (APM: Raven, Court, &

---

[4] We refrained from using a within-participant manipulation because use of the same stimulus materials with different questions can lead to severe carry-over effects concerning strategy use.

[5] Bayes factors for the *t*-tests were computed by means of the BayesFactor R package (Rouder, Speckman, Sun, Morey, & Iverson, 2009) and interpreted according to the classification by Jeffreys (1961).

Raven, 1983; short form: Bors & Stokes, 1998) and verbal subtests from the BIS, respectively.

Participants and Procedure. A total of 93 participants were recruited via posters at the University of Mannheim. The sample consisted of 71 women and 22 men, aged between 18 and 52 years (M = 22.5 years, SD = 5.2). All participants were students (except for three) and were fluent in German.

After providing consent and demographic information, participants completed the two intelligence tests, first the APM and then the BIS subtests, or vice versa. Then, participants were randomly assigned to one of the two conditions described above and worked on the standard RH paradigm relating either to celebrities' age (knowledge condition) or to celebrities' success (neutral condition). Thereby, the sequence of the recognition and the paired-comparison task was randomized across participants. Finally, participants received course credit or a flat fee of 6 Euro, were thanked, and debriefed.

Results

One participant recognized all celebrities and therefore had to be excluded from the following analyses. For the remaining 92 participants, descriptive analyses revealed that fluid intelligence, as captured by the APM, showed strong heterogeneity between participants (M = 7.1, SD = 3.3, Min = 0, Max = 12, and M = 7.6, SD = 2.4, Min = 2, Max = 12, for the knowledge and the neutral condition, respectively, on a scale from 0 to 12). Similarly, crystallized intelligence, as assessed by the verbal BIS tests, varied considerably across participants (M = 44.0, SD = 5.2, Min = 28, Max = 53, and M = 42.1, SD = 6.1 Min = 25, Max = 52, for the knowledge and the neutral condition, respectively, on a scale from 0 to 61). In sum, the results were comparable to previous studies using the same measures (e.g., Bors & Stokes, 1998; Bucik & Neubauer, 1996). Furthermore, as expected, both intelligence measures were correlated within conditions ($\rho(48)$ = .50, p < .001, BF_{10} = 75.04, and $\rho(44)$ = .33, p = .03, BF_{10} = 1.21, for the knowledge and the neutral condition, respectively)[6], replicating previous findings with respect to G_f and G_c (e.g., Cattell, 1963; Cunningham, Clayton, & Overton, 1975).

---

[6] Bayes factors for the Pearson correlations were computed using code provided by Wetzels and Wagenmakers (2012) and interpreted following the classification by Jeffreys (1961).

To analyze our data, we used the hierarchical latent-trait r-model with fluid and crystallized intelligence as predictors of RH-use. This analysis demonstrates that the materials had been chosen in line with our goals: Participants in the knowledge condition showed decisively lower recognition than knowledge validities ($\hat{\mu}^a$ = .52 [.49; .55], $\hat{\mu}^b$ = .77 [.74; .79]; $\hat{\Delta}_{a\text{-}b}$ = -.24 [-.28; -.21]). Thus, this decision environment favored knowledge-use over RH-use. In contrast, recognition and knowledge validities matched almost perfectly in the neutral condition ($\hat{\mu}^a$ = .67 [.64; .69], $\hat{\mu}^b$ = .66 [.63; .69]; $\hat{\Delta}_{a\text{-}b}$ = .01 [-.03; .05]), showing that both strategies were equally efficient here. Replicating the results on RH-use reported above, we found that people adjusted RH-use in line with the relative validity of the recognition cue. Participants in the neutral condition ($\hat{\mu}^r$ = .87 [.84, .91], $\hat{\sigma}^r$ = 0.50 [0.37, 0.66]) used the RH more often than participants in the knowledge condition ($\hat{\mu}^r$ = .06 [.02, .11], $\hat{\sigma}^r$ = 1.27 [0.90, 1.80]; $\hat{\Delta}_r$ = -.81 [-.87, -.73]). Also, there was large variation in RH-use across individuals, as repeatedly found throughout the literature (e.g., Hilbig & Richter, 2011; Newell & Shanks, 2004; Pachur et al., 2008).

In line with our first hypothesis, both intelligence measures exerted negative effects on RH-use in the knowledge condition ($\hat{\beta}^{Fluid}$ = -0.83 [-1.34, -0.37] and $\hat{\beta}^{Cryst}$ = -0.19 [-0.64, 0.24]), indicating that higher levels of intelligence are associated with use of further knowledge. However, as indicated by the BCIs, only the effect of fluid intelligence was significantly different from zero. Thus, whereas fluid intelligence affects RH-use even if individual differences in crystallized intelligence are statistically controlled for, the reverse does not hold. Moreover, in line with our second hypothesis, in the neutral condition we observed that the association between intelligence measures and RH-use was not meaningful, neither for fluid nor for crystallized intelligence ($\hat{\beta}^{Fluid}$ = -0.16 [-0.34, 0.02] and $\hat{\beta}^{Cryst}$ = -0.03 [-0.20, 0.14]). Thus, we conclude that fluid intelligence did not generally influence use versus non-use of the RH; an effect was apparent only under conditions that clearly favor one strategy over the other.

To control for possible confounds between intelligence measures and knowledge about celebrities, we replicated the analyses after additionally including individual recognition and knowledge validities as (standardized) predictors. Inclusion of these control variables affected the regression coefficients for the core predictors of interest only marginally (see Table 1). The effects of individual recognition and knowledge validities showed the expected pattern descriptively: Recognition validity

affected RH-use positively whereas knowledge validity affected it negatively. However, only the effect of knowledge validity in the knowledge condition was reliable. Also, controlling for task order of the recognition and the paired-comparison task did not change the pattern of results. To sum up, we found support for the hypothesis that intelligence moderates adaptive use of the RH and that fluid intelligence is primarily responsible for this effect.

Discussion

Do smarter people make better decisions? To address this question, we investigated the influence of intelligence on individual differences in adaptive use of the recognition heuristic (RH; Goldstein & Gigerenzer, 2002). The RH is one of the simplest and yet surprisingly successful heuristics within the adaptive toolbox (Goldstein & Gigerenzer, 2002). It predicts choice of the recognized object without consideration of additional information. The importance of adaptivity in decision making has been stressed repeatedly (e.g., Simon, 1990). There are numerous studies showing that typical decision makers adapt RH-use according to availability of cognitive resources (e.g., Pohl et al., 2013) and environmental structures (e.g., Hilbig et al., 2010). However, so far, only a single study addressed individual differences in adaptivity by showing that adaptive RH-use declines in older adults (Pachur et al., 2009). As suggested by Mata et al. (2007) and supported by Bröder (2003), this age-related deficit might be caused by a decline in cognitive capacities, especially in fluid and crystallized intelligence. Thus, fluid and crystallized intelligence might be the genuine source of adaptive RH-use, in line with the idea that intelligence reflects the general capacity for successful adaptation (e.g., Neisser et al., 1996).

Our purpose was to test whether intelligence in general, and fluid and crystallized intelligence in particular, affect adaptive RH-use. Therefore, we first reanalyzed a study by Hilbig (2008). In this study, the decision context favored RH-use over knowledge integration, while intelligence was assessed in terms of general intelligence. In a decision context that requires RH-use, we expected RH-use to increase with general intelligence. In addition, we conducted a new experiment with two different decision contexts. The context either disfavored RH-use and thereby fostered use of further knowledge (opposite to the scenario used by Hilbig) or did not favor any of the two strategies. To test which component of general intelligence drives

the effect, intelligence was assessed in terms of fluid and crystallized intelligence, the two major second-order factors of human intelligence (Cattell, 1963). In a decision context that requires knowledge-use, we expected RH-use to decline with intelligence in general and fluid and crystallized intelligence in particular. In contrast, in a context where both strategies are equally advantageous, we expected no influence of either type of intelligence on RH-use.

To analyze our data, we adapted the hierarchical latent-trait r-model (Hilbig et al., 2010; Michalkiewicz & Erdfelder, 2016) based on Klauer's latent-trait approach (2010) to include either general intelligence (for the reanalysis of Hilbig's (2008) data) or fluid and crystallized intelligence (for the analysis of the new experiment) as predictors of RH-use. Thereby we closed a methodological gap in research on individual differences in RH-use. This model allows assessment of the influence of intelligence measures on RH-use in a straightforward way by estimating each individual's probability of RH-use as a linear combination of intelligence measures. Moreover, it features another important advantage: Standard analyses are problematic because they are based on model parameters estimated separately for each participant. This can lead to unreliable and probably biased parameter estimates and to correlation or regression coefficients that are severely underestimated (e.g., Klauer, 2010). In contrast, by applying the hierarchical latent-trait r-model, individual parameters are estimated more reliably by borrowing strength from the imposed group-level structure. Hence, the present study is not only an important step in uncovering *sources* of individual differences in adaptive RH-use but also in improving *methods* for analyzing these individual differences.

The results corroborated our hypotheses. First, our reanalysis of Hilbig's (2008) data showed that general intelligence is positively associated with RH-use in a domain where RH-use is optimal. Second, in the experiment reported herein, fluid and crystallized intelligence were negatively related to RH-use in a domain where knowledge-use is optimal. Whereas the effect of fluid intelligence was reliable when individual differences in crystallized intelligence were statistically controlled for, the reverse did not hold. Thus, crystallized intelligence cannot explain individual differences in RH-use over and above to what can be explained by fluid intelligence alone. In line with corresponding findings of Bröder (2003) and Mata et al. (2007), this suggests that fluid intelligence as measured by the APM is the more fundamental and

powerful predictor of adaptive use of simple decision heuristics. Third, there was no substantial relation between RH-use and either type of intelligence when the decision context did not determine the optimal strategy. Although none of the strategies was superior a priori, it is not surprising that we observed high levels of RH-use here on average. It is plausible that the RH is typically preferred under neutral conditions, simply because it represents the less effortful and less time consuming strategy – the default (Pachur et al., 2009) – compared to knowledge-use. Put differently, because both strategies were equally valid, RH-use might have been preferred on average because the more demanding strategy of knowledge integration was considered not worth the effort (Rieskamp & Hoffrage, 2008).

Note that we observed a weak negative association between intelligence and RH-use even when both strategies (RH-use and knowledge-use) were equally adequate. In other words, in a situation where the context does not uniquely determine the optimal strategy, higher levels of intelligence tend to be associated with a preference for the more complex and cognitively demanding strategy (i.e., knowledge integration). A possible explanation is that intelligent people enjoy cognitively demanding decision strategies. In particular, need for cognition (NFC) reflects the inclination towards effortful cognitive activities (Epstein, Pacini, Denes-Raj, & Heier, 1996) and has been shown to be associated with intelligence (e.g., Fleischhauer et al., 2010). Therefore, it seems plausible that intelligent people prefer complex information integration over simple decision strategies simply because of their higher levels of NFC. Certainly, this idea awaits further investigation.

Despite the converging evidence, our study has some caveats. First, we used a student sample with higher-than-average levels of intelligence. Whereas the APM test was tailored to this type of sample, the BIS tests were designed for the general population. This perhaps explains the restricted range of BIS scores in our sample and might have contributed to the small effect of crystallized intelligence. Second, it is possible that the verbal subtests of the BIS do not capture those aspects of crystallized intelligence that are most closely linked to adaptive strategy selection (cf. Beauducel, Liepmann, Felfe, & Nettelnstroth, 2007). In fact, Bröder (2003) found equivocal results when using the same subtests of the BIS. He found a slightly positive effect of verbal intelligence on use of the TTB heuristic in his main study but not in the corresponding pre-study. Also, Mata et al. (2007) did not find associations between measures of verbal

knowledge and the TTB heuristic. Analogously, one might object that the APM capture selected aspects of fluid intelligence only, possibly leading to an underestimation of the true effect size that could be observed in principle. Third, one might argue that our inference tasks were overly artificial and thus did not invite use of cumulated knowledge acquired in educational contexts or by means of everyday experience. As a consequence, the weak effect of crystallized intelligence does not come as a surprise. Future research should make use of participant samples, inference tasks, and intelligence measures that more closely fit real-world scenarios. Future research should also focus on mechanisms underlying the effect of intelligence on strategy selection. Although we proposed possible mechanisms that might explain the effects of fluid and crystallized intelligence on strategy selection, we did not test the underlying mechanisms directly.

To sum up, our results strongly suggest that intelligence does not affect RH-use in general but rather moderates adaptivity in use of the RH. They thus nicely fit into related lines of research that emphasize the importance of accounting for individual differences in strategy selection (e.g., Newell & Shanks, 2004; Pachur et al., 2008) as well as for potential interaction effects of individual traits and situational influences (e.g., Michalkiewicz & Erdfelder, 2016). In conclusion, this work broadens our understanding of adaptivity as one of the most important factors in decision making by demonstrating individual differences in adaptive strategy selection and revealing one of the cognitive factors underlying these differences.

References

Arnold, N. R., Bayen, U. J., & Böhm, M. F. (2014). Is prospective memory related to depression and anxiety? A hierarchical MPT modelling approach. *Memory, 23,* 1215-1228. doi:10.1080/09658211.2014.969276

Arnold, N. R., Bayen, U. J., & Smith, R. E. (2015). Hierarchical multinomial modeling approaches: An application to prospective memory and working memory. *Experimental Psychology, 62,* 143-152. doi:10.1027/1618-3169/a000287

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57-86. doi:10.3758/BF03210812

Beauducel, A., Liepmann, D., Felfe, J., & Nettelnstroth, W. (2007). The impact of different measurement models for fluid and crystallized intelligence on the correlation with personality traits. *European Journal of Psychological Assessment, 23,* 71-78. doi:10.1027/1015-5759.23.2.71

Bors, D. A., & Stokes, T. L. (1998). Raven's advanced progressive matrices: Norms for first year university students and the development of a short form. *Educational and Psychological Measurement, 58,* 382-398. doi:10.1177/0013164498058003002

Bröder, A. (2003). Decision making with the "Adaptive Toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29*, 611-625. doi:10.1037/0278-7393.29.4.611

Bröder, A., & Eichler, A. (2006). The use of recognition information and additional cues in inferences from memory. *Acta Psychologica, 121*, 275-284. doi:10.1016/j.actpsy.2005.07.001

Bucik, V., & Neubauer, A. C. (1996). Bimodality in the Berlin model of intelligence structure (BIS): A replication study. *Personality and Individual Differences*, *21*, 987-1005. doi:10.1016/S0191-8869(96)00129-8

Castela, M., Kellen, D., Erdfelder, E., & Hilbig, B. E. (2014). The impact of subjective recognition experiences on recognition heuristic use: A multinomial processing tree approach. *Psychonomic Bulletin & Review, 21*, 1131-1138. doi:10.3758/s13423-014-0587-4

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1-22. doi:10.1037/h0046743

Cunningham, W. R., Clayton, V., & Overton, W. (1975). Fluid and crystallized intelligence in young adulthood and old age. *Journal of Gerontology, 30,* 53-55. doi:10.1093/geronj/30.1.53

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology, 71*, 390-405. doi:10.1037/0022-3514.71.2.390

Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models. *Zeitschrift für Psychologie - Journal of Psychology, 217*, 108-124. doi:10.1027/0044-3409.217.3.108

Fleischhauer, M., Enge, S., Brocke, B., Ullrich, J., Strobel, A., & Strobel, A. (2010). Same or different? Clarifying the relationship of need for cognition to personality and intelligence. *Personality and Social Psychology Bulletin, 36*, 82-96. doi:10.1177/0146167209351886

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. Boca Raton, FL: Chapman & Hall.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451-482. doi:10.1146/annurev-psych-120709-145346

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The take the best heuristic. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 75-95). New York, NY: Oxford University Press.

Gigerenzer, G., & Selten, R. (2001). *Bounded rationality: The adaptive toolbox.* London, UK: MIT Press.

Gigerenzer, G., Todd, P. M., & the ABC Research Group (1999). *Simple heuristics that make us smart.* New York, NY: Oxford University Press.

Glöckner, A., & Bröder, A. (2011). Processing of recognition information and additional cues: A model-based analysis of choice, confidence, and response time. *Judgment and Decision Making*, *6*, 23-42. Retrieved from http://journal.sjdm.org/11/rh4/rh4.pdf

Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 37-58). New York, NY: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109*, 75–90. doi:10.1037/0033-295X.109.1.75

Hilbig, B. E. (2008). Individual differences in fast-and-frugal decision making: Neuroticism and the recognition heuristic. *Journal of Research in Personality, 42*, 1641-1645. doi:10.1016/j.jrp.2008.07.001

Hilbig, B. E. (2010). Precise models deserve precise measures: A methodological dissection. *Judgment and Decision Making, 5*, 272-284. Retrieved from http://journal.sjdm.org/10/rh5/rh5.pdf

Hilbig, B. E., Erdfelder, E., & Pohl, R. F. (2010). One-reason decision making unveiled: A measurement model of the recognition heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 123-134. doi:10.1037/a0017518

Hilbig, B. E., Erdfelder, E., & Pohl, R. F. (2012). A matter of time: Antecedents of one-reason decision making based on recognition. *Acta Psychologica, 141*, 9-16. doi:10.1016/j.actpsy.2012.05.006

Hilbig, B. E., Michalkiewicz, M., Castela, M., Pohl, R. F., & Erdfelder, E. (2015). Whatever the cost? Information integration in memory-based inferences depends on cognitive effort. *Memory & Cognition, 43,* 659-671. doi:10.3758/s13421-014-0493-z

Hilbig, B. E., & Pohl, R. F. (2009). Ignorance- versus evidence-based decision making: A decision time analysis of the recognition heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1296-1305. doi:10.1037/a0016565

Hilbig, B. E., Pohl, R. F., & Bröder, A. (2009). Criterion knowledge: A moderator of using the recognition heuristic? *Journal of Behavioral Decision Making, 22,* 510-522. doi:10.1002/bdm.644

Hilbig, B. E., & Richter, T. (2011). Homo heuristicus outnumbered: Comment on Gigerenzer and Brighton (2009). *Topics in Cognitive Science, 3*, 187-196. doi:10.1111/j.1756-8765.2010.01123.x

Horn, J. L. (1991). Measurement of intellectual capabilities: A view of theory. In K. S. McGrew, J. K. Werder, & R. W. Woodcock (Eds.), *Woodcock-Johnson technical manual: A reference on theory and current research* (pp. 197–246). Allen, TX: DLM Teaching Resources.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligence. *Journal of Educational Psychology, 57,* 253-270. doi:10.1037/h0023816

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26,* 107-129. doi:10.1016/0001-6918(67)90011-X

Horn, S. S., Pachur, T., & Mata, R. (2015). How does aging affect recognition-based inference? A hierarchical Bayesian modeling approach. *Acta Psychologica*, *154*, 77-85. doi:10.1016/j.actpsy.2014.11.001

Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test.* Göttingen, Germany: Hogrefe.

Jeffreys, H. (1961). *Theory of Probability* (3rd ed.). Oxford, UK: Oxford University Press.

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika, 75,* 70-98. doi:10.1007/S11336-009-9141-0

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.

Lemaire, P. (2010). Cognitive strategy variations during aging. *Current Directions in Psychological Science, 19,* 363-369. doi:10.1177/0963721410390354

Lunn, D., Spiegelhalter, D., Thomas, A., & Best, N. (2009). The BUGS project: Evolution, critique, and future directions. *Statistics in Medicine, 28,* 3049-3067. doi:10.1002/sim.3680

Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision maker: Cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging, 22,* 796-810. doi:10.1037/0882-7974.22.4.796

Matzke, D., Dolan, C. V., Batchelder, W. H., & Wagenmakers, E.-J. (2015). Bayesian estimation of multinomial processing tree models with heterogeneity in participants and items. *Psychometrika*, *80,* 205-235. doi:10.1007/s11336-013-9374-9

Michalkiewicz, M., & Erdfelder, E. (2016). Individual differences in use of the recognition heuristic are stable across time, choice objects, domains, and presentation formats. *Memory & Cognition.* Advance online publication. doi:10.3758/s13421-015-0567-6

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., … Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77-101. doi:10.1037/0003-066X.51.2.77

Newell, B. R., & Fernandez, D. (2006). On the binary quality of recognition and the inconsequentiality of further knowledge: Two critical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19*, 333-346. doi:10.1002/bdm.531

Newell, B. R., & Shanks, D. R. (2004). On the role of recognition in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 923-935. doi:10.1037/0278-7393.30.4.923

Pachur, T., Bröder, A., & Marewski, J. N. (2008). The recognition heuristic in memory-based inference: Is recognition a non-compensatory cue? *Journal of Behavioral Decision Making, 21*, 183-210. doi:10.1002/bdm.581

Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 983-1002. doi:10.1037/0278-7393.32.5.983

Pachur, T., Mata, R., & Schooler, L. J. (2009). Cognitive aging and the adaptive use of recognition in decision making. *Psychology and Aging, 24*, 901-915. doi:10.1037/a0017211

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 534-552. doi:10.1037/0278-7393.14.3.534

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York, NY: Cambridge University Press.

Pohl, R. F. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19*, 251-271. doi:10.1002/bdm.522

Pohl, R. F., Erdfelder, E., Hilbig, B. E., Liebke, L., & Stahlberg, D. (2013). Effort reduction after self-control depletion: The role of cognitive resources in use of simple heuristics. *Journal of Cognitive Psychology, 25,* 267-276. doi:10.1080/20445911.2012.758101

Pohl, R. F., von Massow, F., & Beckmann, B. (2016). *Developmental differences in using an ecologically valid decision strategy: The case of the recognition heuristic.* Manuscript submitted for publication.

Raven, J. C., Court, J. H., & Raven, J. (1983). *Manual for Raven's progressive matrices and vocabulary scales: Advanced progressive matrices sets I and II.* London, UK: H. K. Lewis.

Richter, T., & Späth, P. (2006). Recognition is used as one cue among others in judgment and decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 150-162. doi:10.1037/0278-7393.32.1.150

Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica, 127,* 258-276. doi:10.1016/j.actpsy.2007.05.004

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16,* 225-237. doi:10.3758/PBR.16.2.225

Scheibehenne, B., & Bröder, A. (2007). Predicting Wimbledon tennis results 2005 by mere player name recognition. *International Journal of Forecasting, 23,* 415-426. doi:10.1016/j.ijforecast.2007.05.006

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129-138. doi:10.1037/h0042769

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41,* 1-19. doi:10.1146/annurev.ps.41.020190.000245

Sternberg, R. J., & Salter W. (1982). Conceptions of intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence.* Cambridge, UK: Cambridge University Press.

Sturtz, S., Ligges, U., & Gelman, A. (2005). R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software*, *12,* 1-16. Retrieved from http://www.jstatsoft.org/v12/i03/paper

Wetzels, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review, 19,* 1057-1064. doi:10.3758/s13423-012-0295-x

Tables

Table 1

Influence of fluid and crystallized intelligence (while controlling for recognition and knowledge validity) on use of the recognition heuristic assessed via the hierarchical latent-trait r-model.

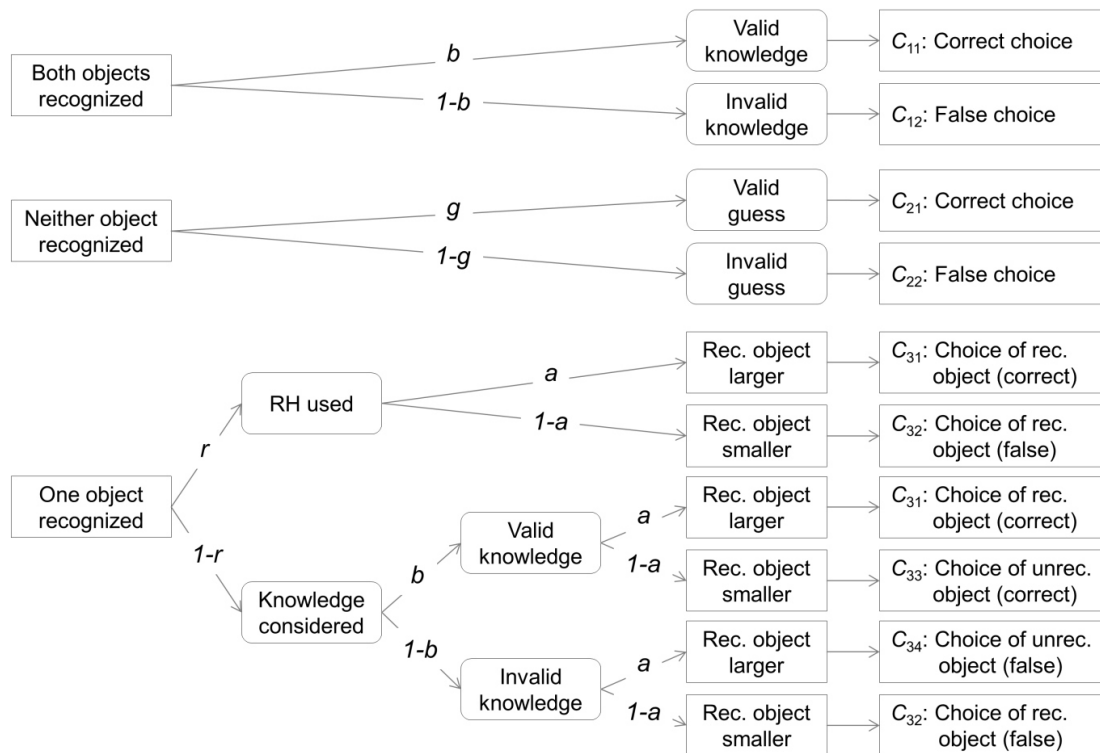|	Regression coefficients (with 95% BCIs)	
Variables	Knowledge condition	Neutral condition
Fluid intelligence (APM)	-0.70 [-1.21, -0.24]	-0.16 [-0.35, 0.03]
Crystallized intelligence (BIS)	-0.26 [-0.70, 0.17]	-0.03 [-0.20, 0.15]
Recognition validity	0.003 [-0.92, 0.92]	0.12 [-0.27, 0.51]
Knowledge validity	-0.62 [-1.23, -0.05]	-0.02 [-0.36, 0.30]

Figures

Figure 1. Illustration of the r-model (adapted from Hilbig et al., 2010). The graph displays the cognitive processes leading to the eight possible choice options C_{ij} ($i \in \{1, 2, 3\}$, $j \in \{1, 2, 3, 4\}$) for the three object pairs (i.e., both, neither, and one object recognized) that can occur in a paired comparison task. Observable events are shown as rectangles, latent states as rectangles with rounded corners. Parameter *r* represents the proportion of RH-use. Parameters *a*, *b*, and *g* represent recognition validity, knowledge validity, and the proportion of correct guessing, respectively.
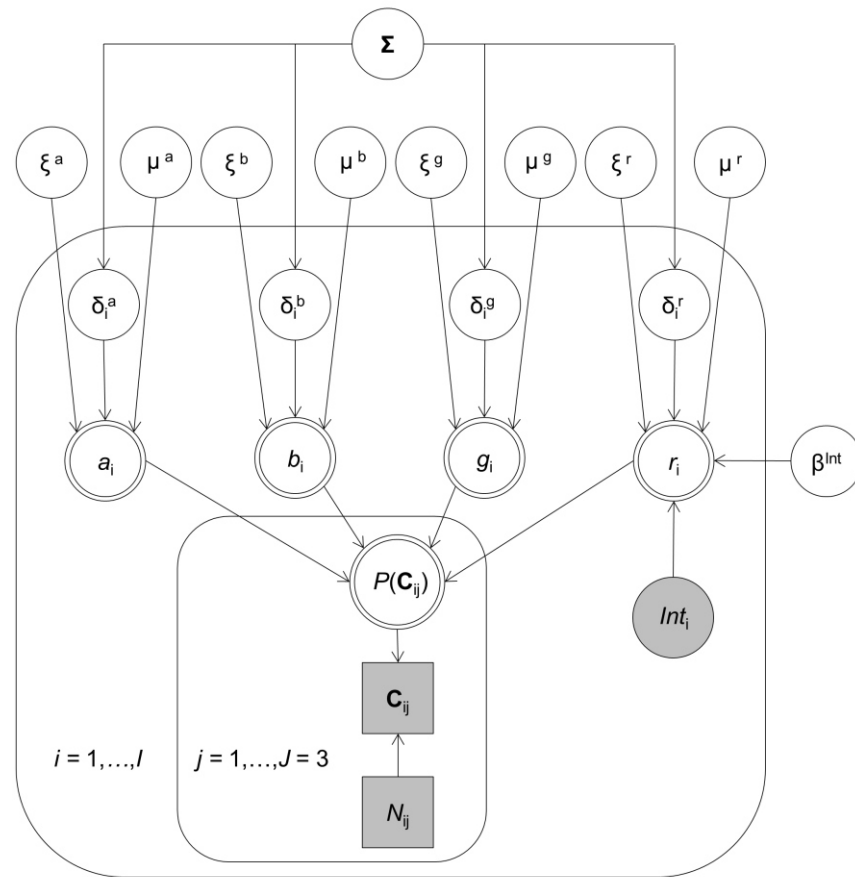
Figure 2. Illustration of the hierarchical r-model including intelligence as a predictor of RH-use (adapted from Michalkiewicz & Erdfelder, 2016). Shaded and unshaded nodes represent observable and unobservable variables, respectively; square and circular nodes represent discrete and continuous variables; single- and double-bordered nodes represent to be estimated and derived variables. The plates display the *I* individuals and the *J* object pairs (i.e., none, one, and both objects recognized). For each individual *i* and each object case *j* the category probabilities $P(\mathbf{C}_{ij})$ are modeled according to the r-model with category counts $\mathbf{C}_{ij} \sim$ Multinomial($P(\mathbf{C}_{ij})$,N_{ij}) and number of observations N_{ij}. Individual model parameters s_i ($s \in \{$ *a*, *b*, *g*, *r* $\}$) are modeled in a probit-transformed parameter space as a linear combination of the group *mean* $\mu^s \sim$ N(0,1), a multiplicative scale parameter $\xi^s \sim$ U(0,100), and individual displacement parameters $\delta_i^s \sim$ MvN(**0**,Σ), with $\Sigma^{-1} \sim$ Wishard(**I**,5). In addition, individual RH-use r_i is modeled as a linear combination of individual scores of intelligence Int_i and a regression coefficient $\beta^{Int} \sim$ N(0,1).

Explaining individual differences in fast-and-frugal decision making:

The impact of need for cognition and faith in intuition

on use of the recognition heuristic

Martha Michalkiewicz, Barbara Minich, and Edgar Erdfelder

University of Mannheim

Author Note

Martha Michalkiewicz, Department of Psychology, University of Mannheim, Schloss, Ehrenhof-Ost, 68131 Mannheim, Germany, email: michalkiewicz@psychologie.uni-mannheim.de; Barbara Minich, email: barbara.minich@hotmail.de; Edgar Erdfelder, Department of Psychology, University of Mannheim, Schloss, Ehrenhof-Ost, 68131 Mannheim, Germany, email: erdfelder@psychologie.uni-mannheim.de.

Please address correspondence to: Martha Michalkiewicz or Edgar Erdfelder, Department of Psychology, School of Social Sciences, University of Mannheim, Schloss, Ehrenhof-Ost, 68131 Mannheim, Germany, phone: +49 621 181 2144, email: michalkiewicz@psychologie.uni-mannheim.de, erdfelder@psychologie.uni-mannheim.de

Abstract

The recognition heuristic (RH) is a decision strategy for paired-comparisons. It predicts choice based on recognition alone without consideration of additional information. Prior work has identified noteworthy individual differences in RH-use, suggesting that individuals have person-specific strategy preferences. To explain these differences, we assessed two plausible personality determinants: Need for cognition (NFC; i.e., enjoyment of cognitively demanding tasks) and faith in intuition (FII; i.e., trust in feelings and impressions). We hypothesized that NFC counteracts RH-use whereas FII fosters it. In our experiment, 82 undergraduates first provided personality measures and then worked on a decision task assessing RH-use in two conditions: A decision context that favored RH-use and thus represented the standard set-up for investigating RH-use, and a neutral context that was expected to boost effects of personality on decision strategies. To test for an effect of NFC and FII on RH-use in either condition, we applied a Bayesian hierarchical multinomial processing tree model that incorporates personality test scores directly into the estimation of RH-use. We found a negative effect of NFC and a positive, yet insignificant, effect of FII in both conditions. Hence, RH-use at least partly reflects a person-specific decision making style as determined by NFC.

Keywords: fast-and-frugal heuristics; individual differences; need for cognition; faith in intuition; hierarchical Bayesian modeling

Explaining individual differences in fast-and-frugal decision making:

The impact of need for cognition and faith in intuition

on use of the recognition heuristic

Introduction

People do not only differ to a great extent in the decisions they make but also in the strategies they use to arrive at these decisions. For instance, when asking several persons to pick the more populous of two cities, some people simply choose the city they recognize (in case they recognize exactly one) whereas others deliberately try to retrieve all their relevant knowledge and integrate it to make a choice. The former strategy is one of the most extensively studied examples of the fast-and-frugal heuristics approach: The recognition heuristic (RH[1]; Goldstein & Gigerenzer, 2002), according to which choice is solely based on recognition. The latter strategy represents a compensatory information-integration approach, that is, integration of knowledge in addition to recognition.

Previous research has revealed substantial individual differences in use of the RH over and above situational influences (e.g., Gigerenzer & Brighton, 2009; Marewski, Gaissmaier, Schooler, Goldstein, & Gigerenzer, 2010; Newell & Shanks, 2004; Pachur, Bröder, & Marewski, 2008). Figure 1 shows a typical example. It illustrates the individual proportions of RH-use in a study by Michalkiewicz and Erdfelder (2016, Exp. 2 Session 1), estimated with the Bayesian hierarchical latent-trait r-model that will be described in detail below. Obviously, even though the decision context is kept constant across participants for these data, RH-use varies significantly. The large variability not only found in this experiment but consistently found in other studies as well gives rise to the idea that people might have person-specific preferences for, or resentments to, certain decision strategies that perhaps can be linked to underlying personality factors. In fact, Michalkiewicz and Erdfelder (2016) showed that RH-use does not vary randomly within and between individuals. Rather, individual RH-use is relatively stable across time and different situational contexts. Thus, it seems likely that individual RH-use reflects a person-specific style of decision making. This opens the way for research on personality determinants of RH-use.

---

[1]Abbreviations: RH = recognition heuristic; NFC = need for cognition; FII = faith in intuition

So far, it has turned out to be difficult to find personality traits that influence strategy selection in fast-and-frugal decision-making (e.g., Bröder, 2012). Specifically, there are only few studies on individual differences in RH-use. On a group level, Pachur, Mata, and Schooler (2009) (for a reanalysis, see Horn, Pachur & Mata, 2015), and Pohl, von Massow, and Beckmann (2016) detected significant age differences in use of the RH. According to their results, RH-use is more frequent in both adolescents and older adults compared to younger adults. Pachur et al. (2009) also tested for associations between measures of inhibitory control and individual RH-use but failed to find significant correlations. On the individual level, Hilbig (2008) investigated the Big Five personality traits and found a positive effect of neuroticism on individual RH-use but no significant effects of agreeableness, conscientiousness, openness, and extraversion. Furthermore, Michalkiewicz, Arden, and Erdfelder (2016) observed that fluid intelligence fostered adaptive use of the RH. Next to these studies, to the best of our knowledge, there is a lack of systematic investigations on personality determinants of RH-use waiting to be filled.

One of the personality theories that allow clear-cut predictions on decision makers' preferences for fast-and-frugal strategies, like the RH, versus cognitively demanding strategies, like knowledge integration, is the Cognitive-Experiential Self-Theory (Epstein, Pacini, Denes-Raj, & Heier, 1996). According to this theory, individuals can draw on two independent systems of information processing: the rational and the experiential system. The rational system is assumed to be analytic, slow, and demanding. It operates through reason, logic, and abstract thought. The experiential system, in contrast, is assumed to be fast, effortless, and associated with affect. Inferences are based, for instance, on concrete examples, categorical thinking, and personal experience. Everybody can make use of both modes of processing, but typically shows preferences for one of them.

These preferences can be measured by two independent personality variables (Epstein et al, 1996): Need for cognition (NFC) and faith in intuition (FII). NFC is characterized as the extent to which people engage in and enjoy cognitively demanding tasks (Cacioppo & Petty, 1982). High NFC is associated with increased elaboration, idea evaluation, and problem solving (Dole & Sinatra, 1998). FII, in contrast, is characterized as the extent to which people trust in intuitive feelings and immediate impressions (Epstein et al, 1996). High FII is associated with affective and associative

information processing that does not rely on verbal reasoning (Zimmerman, Redker, & Gibson, 2011).

Past research has already established some links between NFC and FII on the one hand and heuristics of judgment and decision making on the other hand. For instance, NFC has been found to be associated with less use of the anchoring and adjustment heuristic and partly immunized against both framing and sunk-cost effects in some studies (Carnevale, Inbar, & Lerner, 2011; Epley & Gilovich, 2006). Notably, however, Bröder (2012) failed to find an effect of NFC on use of the Take-the-Best heuristic. By contrast, FII was associated with enhanced use of the ease-of-retrieval heuristic, the representativeness heuristic, the reinforcement heuristic, and liability to framing effects (Alós-Ferrer & Hügelschäfer, 2012; Danziger, Moran, & Rafaely, 2006; Mahoney, Buboltz, Levin, Doverspike, & Svyantek, 2011). Epstein and colleagues (1996, p. 391) even claimed that "heuristic processing represents the natural mode of the experiential system."

In line with these results, Michalkiewicz et al. (2016) recently observed a negative correlation between fluid intelligence and RH-use in a neutral decision context, that is, a context in which simple RH-use and more elaborated knowledge-use were equally effective in terms of accuracy rates. In other words, higher levels of intelligence were weakly associated with a preference for the more complex and cognitively demanding strategy of knowledge integration, despite the fact that the much simpler and less effortful RH resulted in the same proportion of correct decisions. A plausible explanation is that, other things being equal, intelligent people tend to prefer knowledge-use over RH-use because of their higher level of NFC. In fact, a positive correlation between intelligence and NFC was repeatedly reported in the literature (e.g., Fleischhauer et al., 2010; Furnham & Thorne, 2013; Hill et al., 2013; von Stumm, 2013).

Based on the findings outlined above, we hypothesized that (1) NFC affects RH-use negatively. Stated differently, because participants high in NFC enjoy cognitively demanding activities, they will engage in reasoning, and will thus integrate their knowledge more often than participants low in NFC. By contrast, we hypothesized that (2) FII affects RH-use positively. Stated differently, because participants high in FII generally rely more on experiential information, such as recognition, they will use

the RH more often than participants low in FII. In addition, we assessed the Big Five personality traits to control for possible confounds (e.g., Pacini & Epstein, 1999).

To assess whether the hypothesized effects of NFC and FII are moderated by decision strategy effectiveness in a specific decision context, we manipulated recognition and knowledge validities. As outlined by Goldstein and Gigerenzer (2002, p. 78), the recognition validity *a* is defined as the probability of making a correct choice when always following the recognition cue (and thus, using the RH), given that one object is recognized and the other is not. Correspondingly, the knowledge validity *b* is defined as the probability of making a correct choice given that both objects are recognized (and, thus, retrieved knowledge can be used). Hence, the RH is particularly efficient when *a* is high and *b* is low. In contrast, there is no clear advantage of using the RH versus using further knowledge when *a* and *b* are approximately equal.

We created two decision contexts. In the *facilitation condition*, we studied the standard set-up of most RH studies in which use of the RH is more efficient in terms of maximizing correct decisions than use of knowledge (i.e., *a* > *b*). In contrast, in the *neutral condition,* none of the strategies is favored by the decision context (i.e., *a* ≈ *b*). Based on the Cognitive-Experiential Self-Theory, we predicted a negative effect of NFC and a positive effect of FII on RH-use in either condition. However, based on Mischel's (1973) dichotomy of strong versus weak situations, we expected these effects to be less easily detectable in the facilitation condition compared to the neutral condition. With respect to choice behavior, the facilitation condition conforms to what Mischel called a strong situation: It largely determines strategy choice in favor of RH use and thereby presumably masks effects of individual differences in NFC and FII. In contrast, the neutral condition conforms to what Mischel (1973) called a weak situation: It imposes much less constraints on strategy selection. Hence, we expected effects of NFC and FII to be more easily detectable in the neutral condition*,* because situational influences on strategy choice are minimized (cf. Michalkiewicz & Erdfelder, 2016).

Methods

Design & Materials

To assess use of the RH, we relied on the most common paradigm, consisting of a recognition task and a paired-comparison task, presented in random order across participants. In the recognition task, participants provided yes-no recognition judgments

for a set of 25 objects. In the comparison task, participants were asked to decide for a set of 300 pairs (resulting from pairing the 25 objects exhaustively) which of two objects had a higher value with respect to the criterion of interest.

We tested our hypotheses in two conditions manipulated within participants. In the facilitation condition, we used the world's most successful musicians, defined by the number of records sold worldwide. Participants provided recognition judgments for a set of 25 musicians randomly drawn from the world's 150 most successful musicians and compared them with respect to their success (i.e., records sold worldwide). In this condition, RH-use is more efficient than knowledge-use on a-priori grounds since the recognition validity is significantly higher than knowledge validity ($a = .69$, $SD = .14$, and $b = .60$, $SD = .08$; $t(63) = 4.17$, $p < .001$, $BF_{10} = 224.77$; cf. Michalkiewicz & Erdfelder, 2016)[2].

In the neutral condition, by contrast, we used the world's longest rivers. Participants provided recognition judgments for a set of rivers randomly drawn from the world's 60 longest rivers and compared them with respect to their lengths. In this condition, both RH-use and knowledge-use are equally efficient on a-priori grounds: According to an independent prestudy, recognition and knowledge validities are approximately equal for this decision domain ($a = .65$, $SD = .10$, and $b = .62$, $SD = .21$; $t(21) = 0.53$, $p = .60$, $BF_{10} = 0.25$).

Participants & Procedure

A total of 82 participants (38 females), aged between 16 and 39 years ($M = 22.6$, $SD = 3.2$) were recruited at the University of Mannheim. All participants were students (except for six) and spoke German fluently (except for one).

After providing consent and demographic information, participants first completed two personality questionnaires: The Rational-Experiential Inventory to access NFC and FII (Keller, Bohner, & Erb, 2000) and the NEO FFI to control for an influence of the Big Five personality traits (Borkenau & Ostendorf, 1989). Then, participants worked on the decision task measuring RH-use in the two conditions described above, arranged in a random order per participant. Finally, participants received a flat fee of 3 Euro, were thanked and debriefed.

---

[2] We computed Bayes factors BF_{10} for the t-tests using the BayesFactor R package (Rouder, Speckman, Sun, Morey, & Iverson, 2009) and interpreted them according to Jeffreys (1961).

Model

To analyze our data, we relied on the r-model (Hilbig, Erdfelder, & Pohl, 2010), a multinomial processing tree (MPT) model. The r-model (see Figure 2) illustrates the cognitive processing paths that can occur for the three possible cases in a paired-comparison task: Both, one, or none of the two objects is recognized. Responses provided in the paired-comparison task are assigned to eight mutually exclusive categories. The observed frequencies for these categories are explained in terms of four latent parameters, namely, recognition validity (parameter a), knowledge validity (parameter b), probability of correct guessing (parameter g), and, most importantly, probability of RH-use (parameter r). Model parameters are typically estimated using the expectation-maximization algorithm of maximum likelihood estimation (Hu & Batchelder, 1994), assuming a (homogeneous) joint multinomial distribution of the data. For a comprehensive introduction and a review on MPT models, see Batchelder and Riefer (1999) and Erdfelder et al. (2009).

In contrast to ad-hoc measures of RH-use, the r-model provides a measure that can be directly interpreted as the probability of RH-use free of confounds with knowledge integration and other possible influences (for a comparison of different RH measures, see Hilbig, 2010). Specifically, the r parameter represents the proportion of pure RH-use whereas the complementary probability $1 - r$ represents the proportion of using further knowledge retrieved from memory.

The version of the r-model illustrated in Figure 2 assumes fixed parameters and thus cannot account for individual differences in parameter values. Based on Klauer's (2010) latent-trait approach to MPT models, we thus generalized the r-model to a hierarchical version (see Figure 3; cf. Michalkiewicz & Erdfelder, 2016). The basic idea is to specify the model on two levels: On the individual level, a separate set of parameters (r_i, a_i, b_i, g_i) is defined for each participant i, $i = 1, \ldots, N$, according to the r-model described above. On the hierarchical group level, the probits of these individual parameters are assumed to follow a multivariate normal distribution that captures the variability between individuals using group-level parameters (i.e., the mean vector (μ^a, μ^b, μ^g, μ^r) and the variance-covariance matrix Σ of the probit-transformed parameters). Furthermore, to assess the influence of NFC and FII on RH-use, personality test scores are incorporated directly into the estimation of RH-use r_i in terms of a probit regression model.

More precisely, we modeled the probit-transformed individual RH-use r_i as a linear combination of the group-level mean of RH-use μ^r, a multiplicative scale parameter ξ^r, individual deviations from the group mean δ_i^r, and, additionally, z-transformed test scores in NFC and FII along with their regression parameters β^{NFC} and β^{FII}, respectively. Thus, our core model equation was $\phi^{-1}(r_i) = \mu^r + \xi^r \cdot \delta_i^r + \beta^{NFC} \cdot NFC_i + \beta^{FII} \cdot FII_i$, where $\phi^{-1}(.)$ denotes the inverse of the standard normal distribution function.

The hierarchical latent-trait r-model has a number of advantages. Consider the standard method to relate individual RH-use to personality variables: Model parameters are estimated separately for each participant and then regressed on or correlated with personality test scores. This way of analyzing data can be problematic for several reasons. First, fitting the r-model separately to each individual's data can lead to unreliable and biased parameter estimates because of the relatively small number of observations per participant (e.g., Hilbig Erdfelder, & Pohl, 2010). Second, if the parameter estimates are subject to measurement noise, the relation to external variables will in general be severely underestimated (e.g., Spearman, 1904). In contrast, the hierarchical latent-trait r-model estimates individual-level parameters more reliably by making use of the hierarchical group-level structure. Furthermore, it allows assessment of personality influences on RH-use in a single step by incorporating (z-transformed) test scores of NFC and FII into the estimation of individual RH-use r_i. The estimated regression coefficients β^{NFC} and β^{FII} are thus automatically adjusted for the uncertainty in the individual parameter estimates of RH-use.

All analyses were conducted within the Bayesian framework using Markov chain Monte Carlo sampling techniques by means of OpenBUGS (Lunn, Spiegelhalter, Thomas, & Best, 2009) and R2WinBUGS (Sturtz, Ligges, & Gelman, 2005). With this method, prior beliefs represented by prior distributions (being either informative or rather vague) are updated by the observed data, resulting in posterior distributions. Properties of the posterior distributions are used to summarize the results. Specifically, the mean can be interpreted as a point estimate while the 95% Bayesian credible interval (BCI) quantifies its precision. For a comprehensive introduction to hierarchical MPT-models and Bayesian modeling, see for instance Lee and Wagenmakers (2013).

In our analyses, we defined priors following the example of Matzke, Dolan, Batchelder, and Wagenmakers (2015). For each analysis, we ran three chains with

500,000 iterations, a thinning rate of 10, and a burn-in period of 100,000. For all parameter estimates, we ensured chain convergence (all $\hat{R}$ < 1.01; Gelman, Carlin, Stern, & Rubin, 2004) and sufficiently large effective samples (Kruschke, 2014). In the subsequent section, we report the means of the posterior distributions along with their 95% BCIs for all parameters of interest, namely $\hat{\mu}$, $\hat{\sigma}$ (derived from the variance-covariance matrix $\boldsymbol{\Sigma}$), $\hat{\beta}^{NFC}$, and $\hat{\beta}^{FII}$. Here, $\hat{\mu}^s$ represents the group-level mean and $\hat{\sigma}^s$ the group-level standard deviation ($s \in \{a, b, r\}$) for recognition validity a, knowledge validity b, and RH-use r, respectively. Most importantly, the regression coefficients $\hat{\beta}^{NFC}$ and $\hat{\beta}^{FII}$ represent the influence of NFC and FII on RH-use, respectively.

Results

Six participants had to be excluded from the analyses: One participant indicated insufficient language skills, four participants recognized all musicians, and one did not recognize any river. For the remaining 76 participants, we applied the hierarchical latent-trait r-model described in the previous section. As expected and intended, mean recognition validity was higher than mean knowledge validity in the facilitation condition ($\hat{\mu}^a$ = .73 [.70, .76], $\hat{\sigma}^a$ = .32 [.26, .39], $\hat{\mu}^b$ = .64 [.62, .66], $\hat{\sigma}^b$ = .22 [.18, .27]; $\hat{\Delta}_{a-b}$ = .09 [.05, .13]), whereas both validities were almost identical in the neutral condition ($\hat{\mu}^a$ = .58 [.56, .60], $\hat{\sigma}^a$ = .19 [.15, .24], $\hat{\mu}^b$ = .58 [.56, .60], $\hat{\sigma}^b$ = .18 [.14, .22]; $\hat{\Delta}_{a-b}$ = .005 [-.02, .03]). Replicating previous studies, we identified substantial individual differences in RH-use as indicated by large standard deviations of r ($\hat{\mu}^r$ = .76 [.70, .81], $\hat{\sigma}^r$ = .70 [.57, .85], and $\hat{\mu}^r$ = .58 [.50, .65], $\hat{\sigma}^r$ = .82 [.68, 1.00], for the facilitation and the neutral condition, respectively). Furthermore, NFC and FII showed considerable variability across individuals (NFC: M = 5.05, SD = 0.94; FII: M = 4.38, SD = 0.79), and were almost uncorrelated ($\rho(76)$ = .07, p = .58, BF_{10} = 0.11)[3], as hypothesized by Epstein et al. (1996).

In line with our first hypothesis, we found negative effects of NFC on RH-use in both conditions ($\hat{\beta}^{NFC}$ = -.17 [-.32, -.01] and $\hat{\beta}^{NFC}$ = -.20 [-.38, -.03] for the facilitation and the neutral condition, respectively), with a slightly larger effect in the neutral condition. In line with our second hypothesis, we found weak positive effects of FII on RH-use ($\hat{\beta}^{FII}$ = .01 [-.14, .17] and $\hat{\beta}^{FII}$ = .05 [-.12, .22] for the facilitation and the neutral

---

[3] The Bayes factor for the correlation was computed according to Wetzels and Wagenmakers (2012).

condition, respectively). However, both 95% BCIs include zero and thus indicate that the influence of FII was not reliable in either condition. Inclusion of the Big Five personality traits as additional (z-transformed) predictors into the model did not change the pattern of results (see Table A1 in the Appendix). Hence, NFC affects RH-use even if individual differences in the Big Five traits are statistically controlled for. To check for further possible confounds, we replicated the analyses while controlling for (z-transformed) individual recognition and knowledge validities in addition. Again, this changed the results only marginally (see Table A2 in the Appendix). As one might expect, individual recognition validities affected RH-use positively (i.e., the higher the individual RH success rate, the more often the RH is used) whereas individual knowledge validities affect RH-use negatively (i.e., the higher the individual knowledge-use success rate, the less RH use). However, in contrast to the effect of NFC, both validity effects were insignificant.

Discussion

People differ in their preferences for specific decision strategies. Some individuals seem to prefer fast and simple strategies, like the RH (Goldstein & Gigerenzer, 2002), while others seem to favor more complex and cognitively demanding strategies, such as knowledge integration. Since these individual differences are stable across time, sets of choice objects, presentation modalities, and even decision domains (Michalkiewicz & Erdfelder, 2016), we tested a personality framework – the Cognitive-Experiential Self-Theory – as a potential explanation (Epstein et al., 1996). According to Epstein and colleagues, individuals may rely on two systems of information processing, the rational and the experiential system, that are independent and can work simultaneously. The rational system is assumed to be analytic and effortful, while the experiential system is assumed to be automatic and effortless. Epstein and colleagues argued that most individuals have a preference for one of the two systems. These preferences correspond to two personality traits: NFC (i.e., enjoyment of cognitively demanding activities) and FII (i.e., trust in feelings and impressions). Based on the Cognitive-Experiential Self-Theory, we predicted NFC to diminish RH-use and FII to foster it. In addition, based on Mischel's (1973) dichotomy of strong versus weak situations, we expected relatively small effects of NFC and FII when the optimal decision strategy is largely determined by the decision context

because all people adapt their decision strategies to situational constraints to some degree. In contrast, we expected more pronounced effects of NFC and FII when the decision context does not favor any strategy because contextual influences on strategy selection are minimized.

We analyzed our data with a hierarchical latent-trait version of the r-model (Hilbig, Erdfelder, & Pohl, 2010; Klauer, 2010; Michalkiewicz & Erdfelder, 2016), including NFC and FII as predictors of RH-use. Using this formal framework, the effect of NFC and FII can be estimated more reliably than by means of standard analyses (i.e., applying the r-model to the data of each participant separately followed by a regression of RH-use on personality measures). In line with our first hypothesis, NFC affected RH-use negatively, irrespective of whether the decision domain fostered RH-use or not. In other words, participants who enjoy engagement in cognitively demanding tasks preferred knowledge-use over RH-use more than those who dislike demanding tasks. This is consistent with prior work showing that people high in NFC make less use of certain heuristics (e.g., Epley & Gilovich, 2006) and not only incorporate more information overall but also more varied information (Cacioppo & Petty, 1982; Nair & Ramnarayan, 2000). Notably, the observed difference of NFC effects on RH-use between decision domains was only marginal. Hence, NFC can be considered a rather general predictor of strategy selection, because the negative influence of NFC on RH-use manifests itself even under strong situational demands that foster RH use, and not only under neutral conditions where contextual influences are minimized.

When assessing the strength of NFC effects on RH-use, it is important to keep in mind that Michalkiewicz and Erdfelder (2016) observed test-retest correlations for RH-use ranging between .33 and .80 (depending on the length of the test-retest interval and differences in decision objects, domains, and presentation modalities). Since validity coefficients cannot be expected to exceed test-retest correlations, any correlation between NFC and RH-use close to .30 can be considered evidence for a strong effect of NFC. By re-transforming the NFC regression coefficients reported in the results section into bivariate correlations, we derived correlation estimates of $\hat{\rho} = -.23$ [-.43, -.02] and $\hat{\rho} = -.24$ [-.42, -.04] for the facilitation and the neutral condition, respectively. Thus, compared to the results of Michalkiewicz and Erdfelder (2016), the effects observed in the present study can be classified as moderate to strong.

It is also interesting to compare our results to findings of Hilbig, Scholl, and Pohl (2010). At a first glance, our results seem to contradict theirs. Hilbig and colleagues found that participants who were prompted to think deliberately applied the RH more often compared to those who were prompted to think intuitively (still, both groups applied the RH roughly equally often as in previous experiments). To explain this finding, they suggested that a deliberate mode of processing is more effortful and demanding, and that people use the RH more often to reduce this effort. By contrast, they suggested that an intuitive mode of processing is more automatic and effortless, thus rendering effort reduction less necessary.

In our study, we apparently found the opposite result, namely, that participants high in NFC (i.e., those preferring a deliberate processing style) used the RH less often. However, there is one important difference between these two studies: We assessed effects of an individual trait whereas Hilbig, Scholl, and Pohl (2010) prompted their subjects to use a specific mode of thinking irrespective of their individual disposition. One way to reconcile the conflicting results is based on the following argument: In the experiment by Hilbig and colleagues, individuals who disfavored deliberate processing (i.e., low NFC individuals) but were instructed to use it in Hilbig, Scholl, and Pohl's (2010) experiment may have been overwhelmed by the cognitive effort associated with search for, reflection of, and integration of relevant further knowledge (Cacioppo & Petty, 1982). They probably compensated this cognitive overload by using less effortful strategies (i.e., the RH) even more frequently than when instructed to decide in line with their processing style. In contrast, individuals high in NFC think deliberately by default and thus can handle the cognitive effort associated with information search and integration (Cacioppo & Petty, 1982). They enjoy cognitively demanding tasks and therefore probably avoid using the RH even when instructed to decide more intuitively and automatically. Hence, the result observed by Hilbig and colleagues might be due to a subsample of participants low in NFC. Unfortunately, since measures of NFC or FII were not obtained in their study, there is no direct way to test this explanation.

Unexpectedly, we did not find a noticeable effect of FII on RH-use, not even in the neutral condition where it should be easiest to observe (Michalkiewicz & Erdfelder, 2016). In other words, participants who principally trust their feelings and generally prefer experiential cues did not show significantly enhanced use of the RH. One possible explanation might be that the vast majority of participants were students, a

group of persons typically high in NFC. In line with this explanation, the average NFC value was quite high in our study whereas the average FII value was low compared to previous studies (e.g., Epstein et al., 1996; Pacini & Epstein, 1999). Hence, NFC effects were potentially masking FII effects. Furthermore, FII scores did not exhaust the possible scale range in our student sample (*Min* = 2.53, *Max* = 6.47, on a 1 to 7 scale). A replication study using a non-student sample might shed more light on this problem.

Also note that our null finding concerning FII does not seem to be particularly uncommon. Work by Keller et al. (2000), for example, has already shown that high FII does not automatically lead to more heuristic use. This suggests that the causal link between the experiential information processing trait (as indexed by the FII scale) and preferential use of simple heuristics is less clear than presumed by Epstein and colleagues (1996).

In sum, we found supportive evidence for the hypothesis that NFC hampers RH-use. In contrast, we found no convincing evidence for the corresponding hypothesis that FII boosts RH-use. Thus, our work contributes to clarifying individual differences in heuristic use and to identifying their sources (e.g., Pachur et al., 2008). An important goal for future research will be to promote this research program further (e.g., by focusing on potential moderators for FII effects on RH-use) and perhaps extend it to other potential personality determinants of decision strategy selection. To conclude, our current work suggests that this line of research is likely to advance our knowledge on determinants of decision strategy choice and thus to add to a comprehensive theory of decision making.

Declaration of conflict of interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

References

Alós-Ferrer, C., & Hügelschäfer, S. (2012). Faith in intuition and behavioral biases. *Journal of Economic Behavior & Organization, 84,* 182-192. doi:10.1016/j.jebo.2012.08.004

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57-86. doi:10.3758/BF03210812

Borkenau, P., & Ostendorf, F. (1989). Untersuchungen zum Fünf-Faktoren-Modell der Persönlichkeit und seiner diagnostischen Erfassung. *Zeitschrift für Differentielle und Diagnostische Psychologie, 10,* 239-251. Retrieved from http://pub.uni-bielefeld.de/luur/download?func=downloadFile&recordOId=1780810&fileOId=2313037

Bröder, A. (2012). The quest for take the best – Insights and outlooks from experimental research. In P. M. Todd, G. Gigerenzer, & the ABC Research Group (Eds.), *Ecological rationality: Intelligence in the world* (pp. 216-240). New York, NY: Oxford University Press.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42,* 116-131. doi:10.1037/0022-3514.42.1.116

Carnevale, J. J., Inbar, Y., & Lerner, J. S. (2011). Individual differences in need for cognition and decision-making competence among leaders. *Personality and Individual Differences, 51,* 274-278. doi:10.1016/j.paid.2010.07.002

Danziger, S., Moran, S., & Rafaely, V. (2006). The influence of ease of retrieval on judgment as a function of attention to subjective experience. *Journal of Consumer Psychology, 16,* 191-195. doi:10.1207/s15327663jcp1602_9

Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist, 33,* 109-128. doi:10.1080/00461520.1998.9653294

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science, 17,* 311-318. doi:10.1111/j.1467-9280.2006.01704.x

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of*

Personality and Social Psychology, 71, 390-405. doi:10.1037/0022-3514.71.2.390

Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models. *Zeitschrift für Psychologie - Journal of Psychology, 217,* 108-124. doi:10.1027/0044-3409.217.3.108

Fleischhauer, M., Enge, S., Brocke, B., Ullrich, J., Strobel, A., & Strobel, A. (2010). Same or different? Clarifying the relationship of need for cognition to personality and intelligence. *Personality and Social Psychology Bulletin, 36,* 82-96. doi:10.1177/0146167209351886

Furnham, A., & Thorne, J. D. (2013). Need for cognition: Its dimensionality and personality and intelligence correlates. *Journal of Individual Differences*, *34*, 230-240. doi:10.1027/1614-0001/a000119

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.)*.* London, UK: Chapman and Hall.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*, 107-143. doi:10.1111/j.1756-8765.2008.01006.x

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*, 75-90. doi:10.1037/0033-295X.109.1.75

Hilbig, B. E. (2008). Individual differences in fast-and-frugal decision making: Neuroticism and the recognition heuristic. *Journal of Research in Personality, 42*, 1641-1645. doi:10.1016/j.jrp.2008.07.001

Hilbig, B. E. (2010). Precise models deserve precise measures: A methodological dissection. *Judgment and Decision Making, 5*, 272-284. Retrieved from http://journal.sjdm.org/10/rh5/rh5.pdf

Hilbig, B. E., Erdfelder, E., & Pohl, R. F. (2010). One-reason decision making unveiled: A measurement model of the recognition heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 123-134. doi:10.1037/a0017518

Hilbig, B. E., Scholl, S. G., & Pohl, R. F. (2010). Think or blink − is the recognition heuristic an "intuitive" strategy? *Judgment and Decision Making, 5*, 300-309. Retrieved from http://journal.sjdm.org/10/rh6/rh6.pdf

Hill, B. D., Foster, J. D., Elliott, E. M., Shelton, J. T., McCain, J., & Gouvier, W. D. (2013). Need for cognition is related to higher general intelligence, fluid intelligence, and crystallized intelligence, but not working memory. *Journal of Research in Personality*, *47*, 22-25. doi:10.1016/j.jrp.2012.11.001

Horn, S. S., Pachur, T., & Mata, R. (2015). How does aging affect recognition-based inference? A hierarchical Bayesian modeling approach. *Acta Psychologica*, *154*, 77-85. doi:10.1016/j.actpsy.2014.11.001

Hu, X., & Batchelder, W. H. (1994). The statistical analysis of general processing tree models with the EM algorithm. *Psychometrika*, *59*, 21-47. doi:10.1007/BF02294263

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Keller, J., Bohner, G., & Erb, H. (2000). Intuitive und heuristische Urteilsbildung – verschiedene Prozesse? Präsentation einer deutschen Fassung des 'Rational-Experiential Inventory' sowie neuer Selbstberichtskalen zur Heuristiknutzung. *Zeitschrift für Sozialpsychologie, 31,* 87-101. doi:10.1024//0044-3514.31.2.87

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika, 75,* 70-98. doi:10.1007/S11336-009-9141-0

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course.* New York, NY: Cambridge University Press.

Lunn, D., Spiegelhalter, D., Thomas, A., & Best, N. (2009). The BUGS project: Evolution, critique and future directions (with discussion). *Statistics in Medicine, 28*, 3049-3082. doi:10.1002/sim.3680

Mahoney, K. T., Buboltz, W., Levin, I. P., Doverspike, D., & Svyantek, D. J. (2011). Individual differences in a within-subjects risky-choice framing study. *Personality and Individual Differences, 51*, 248-257. doi:10.1016/j.paid.2010.03.035

Marewski, J. N., Gaissmaier, W., Schooler, L. J., Goldstein, D. G., & Gigerenzer, G. (2010). From recognition to decisions: Extending and testing recognition-based models for multialternative inference. *Psychonomic Bulletin & Review, 17*, 287-309. doi:10.3758/PBR.17.3.287

Matzke, D., Dolan, C. V., Batchelder, W. H. & Wagenmakers, E.-J. (2015). Bayesian estimation of multinomial processing tree models with heterogeneity in participants and items. *Psychometrika, 80*, 205-235. doi:10.1007/s11336-013-9374-9

Michalkiewicz, M., & Erdfelder, E. (2016). Individual differences in use of the recognition heuristic are stable across time, choice objects, domains, and presentation formats. *Memory & Cognition*. Advance online publication. doi:10.3758/s13421-015-0567-6

Michalkiewicz, M., Arden, K., & Erdfelder, E. (2016). *Do smarter people make better decisions? The influence of intelligence on adaptive use of the recognition heuristic.* Manuscript submitted for publication

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, *80*, 252-283. doi:10.1037/h0035002

Nair, K. U., & Ramnarayan, S. (2000). Individual differences in need for cognition and complex problem solving. *Journal of Research in Personality, 34,* 305-328. doi:10.1006/jrpe.1999.2274.

Newell, B. R., & Shanks, D. R. (2004). On the role of recognition in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 923-935. doi:10.1037/0278-7393.30.4.923

Pachur, T., Bröder, A., & Marewski, J. N. (2008). The recognition heuristic in memory-based inference: Is recognition a non-compensatory cue? *Journal of Behavioral Decision Making, 21*, 183-210. doi:10.1002/bdm.581

Pachur, T., Mata, R., & Schooler, L. J. (2009). Cognitive aging and the adaptive use of recognition in decision making. *Psychology and Aging, 24*, 901-915. doi:10.1037/a0017211

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology, 76,* 972-987. doi:10.1037/0022-3514.76.6.972

Pohl, R. F., von Massow, F., & Beckmann B. (2016). *Developmental differences in using an ecologically valid decision strategy: The case of the recognition heuristic.* Manuscript submitted for publication

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16,* 225-237. doi:10.3758/PBR.16.2.225

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology, 15,* 72-101. doi:10.2307/1412159

Sturtz, S., Ligges, U., & Gelman, A. (2005). R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software, 12,* 1-16. Retrieved from http://www.jstatsoft.org/v12/i03/paper

von Stumm, S. (2013). Investment traits and intelligence in adulthood. Assessment and associations. *Journal of Individual Differences, 34,* 82-89. doi:10.1027/1614-0001/a000101

Wetzels, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review, 19,* 1057-1064. doi:10.3758/s13423-012-0295-x

Zimmerman, I., Redker, C., & Gibson, B. (2011). The role of faith in intuition, need for cognition and method of attitude formation in implicit-explicit brand attitude relationship strength. *Journal of Consumer Psychology, 21,* 290-301. doi:10.1016/j.jcps.2011.04.001
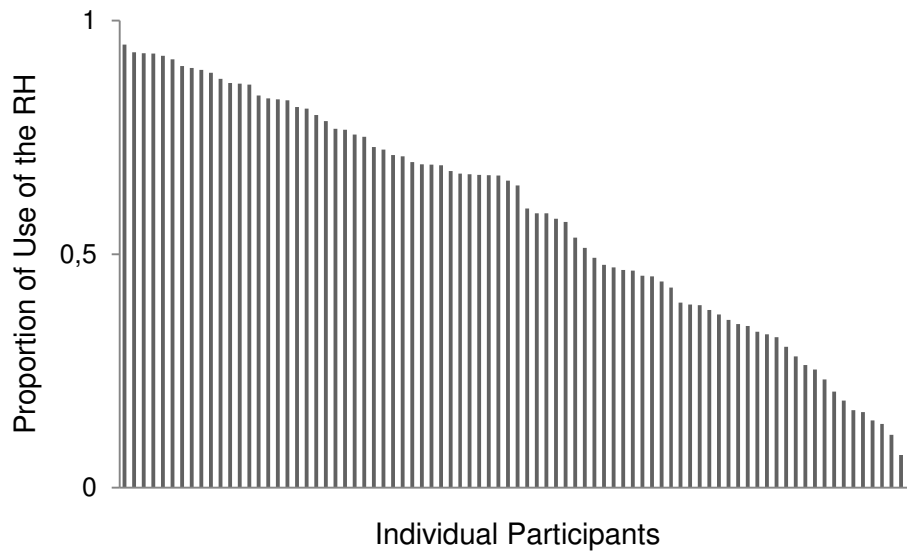
IMPACT OF NFC AND FII ON RH

Figures

Figure 1. Proportions of use of the RH per participant for the data from Michalkiewicz and Erdfelder (2016, Exp. 2 Session 1). Individual RH-use is estimated via the *r* parameter of the hierarchical latent-trait r-model described in the methods section, and ordered by size.
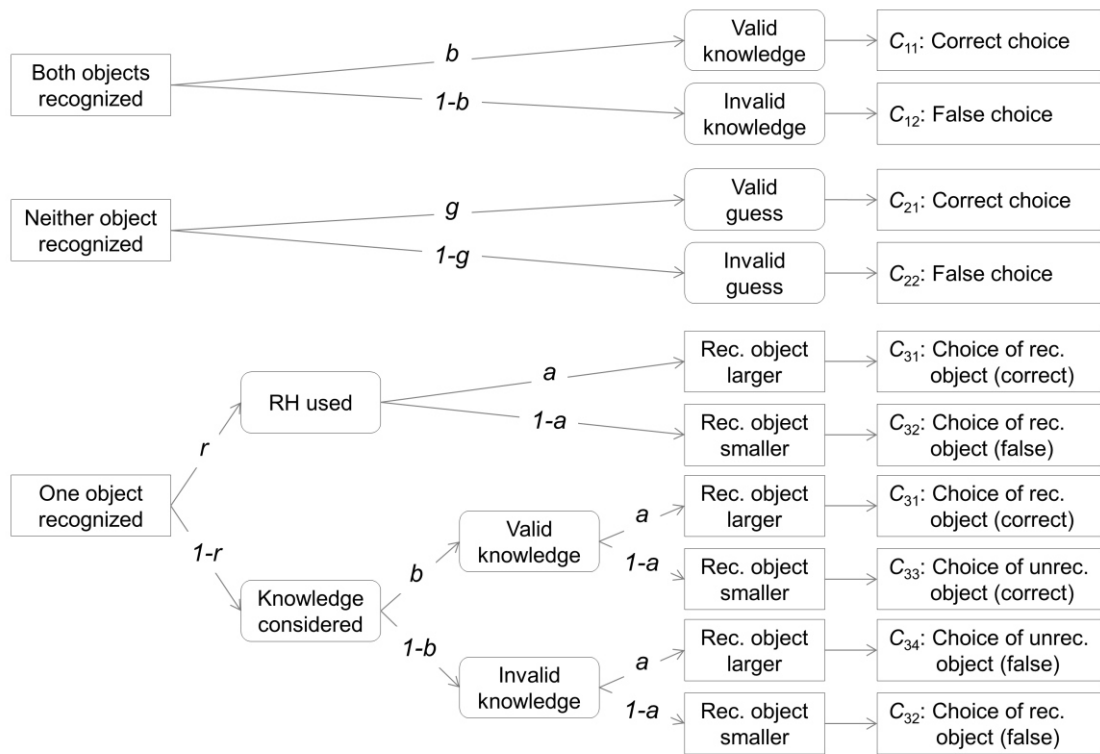
Figure 2. Illustration of the r-model (Hilbig, Erdfelder, & Pohl, 2010). Observable and unobservable events are presented as rectangles vs. rectangles with rounded corners. For each of the $J = 3$ object pairs (both, none, or one object recognized), responses are assigned to one of $m \in \{1,..,M\}$ categories C_{jm} and modeled by means of parameters a, b, g, and r representing recognition validity, knowledge validity, probability of correct guessing, and probability of RH-use, respectively.
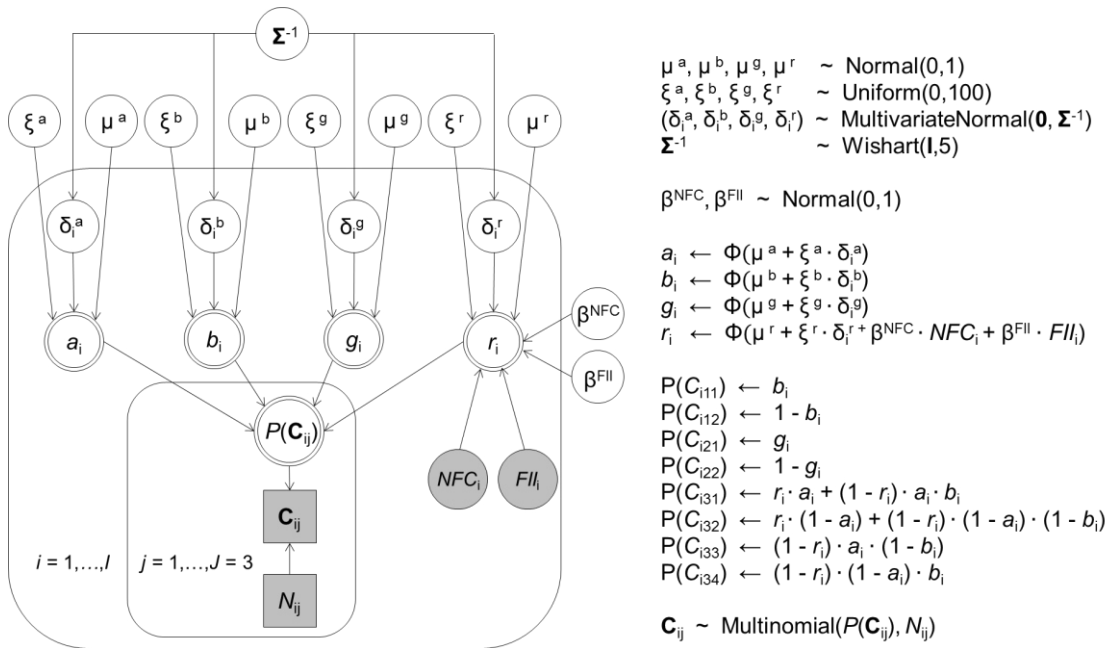
Figure 3. Illustration of the hierarchical latent-trait r-model including NFC and FII as predictors of RH-use (adapted from Michalkiewicz & Erdfelder, 2016). The graph shows the relations among observed data and latent model parameters, presented as shaded and unshaded notes, respectively. Discrete and continuous variables are presented as square vs. circular nodes, to be estimated and derived variables as single- vs. double-bordered nodes. The plates show replications over I individuals and the J object cases (none, one, or both objects recognized). $P(\mathbf{C}_{ij})$ represents the category probabilities of the r-model for vectors of category counts $\mathbf{C}_{ij}$ and N_{ij} observations for individual i and object case j. For individual parameters s_i ($s \in \{a, b, g, r\}$) μ^s depicts the group mean, δ_i^s the individual deviation from the group mean, and ξ^s a multiplicative scale parameter. The influence of NFC and FII on RH-use r_i is assessed by regression parameters β^{NFC} and β^{FII}.

Appendix

Table A1.

Influence of need for cognition, faith in intuition, and the Big Five personality traits on use of the recognition heuristic assessed via the hierarchical latent-trait r-model.

Variables	Regression coefficients (with 95% BCIs)	
	Facilitation condition	Neutral condition
Need for cognition	-.27 [-.47, -.06]	-0.32 [-0.56, -0.07]
Faith in intuition	.03 [-.13, .20]	0.07 [-0.13, 0.26]
Neuroticism	.04 [-.13, .22]	-0.02 [-0.23, 0.18]
Extraversion	.03 [-.14, .20]	-0.07 [-0.26, 0.13]
Openness for new experiences	.21 [.01, .41]	0.17 [-0.06, 0.41]
Agreeableness	-.01 [-.17, .14]	0.03 [-0.16, 0.21]
Conscientiousness	-.10 [-.26, .20]	0.09 [-0.10, 0.28]

Note. Estimates are based on one million iterations per chain.

Table A2.

Influence of need for cognition, faith in intuition, recognition and knowledge validities
on use of the recognition heuristic assessed via the hierarchical latent-trait r-model.

Variables	Regression coefficients (with 95% BCIs)	
	Facilitation condition	Neutral condition
Need for cognition	-0.17 [-0.33, -0.01]	-.21 [-.39, -.03]
Faith in intuition	0.02 [-0.14, 0.17]	.05 [-.12, .22]
Recognition validity	0.18 [-0.33, 0.68]	.08 [-.46, .61]
Knowledge validity	-0.06 [-0.45, 0.31]	.03 [-.21, .27]

Note. Estimates are based on one million iterations per chain.