Judgment and Decision Making, Vol. 6, No. 1, February 2011, pp. 7–22

Threshold models of recognition and the recognition heuristic

Edgar Erdfelder* Carolina E. Küpper-Tetzel† Sandra D. Mattern†

Abstract

According to the recognition heuristic (RH) theory, decisions follow the recognition principle: Given a high validity of the recognition cue, people should prefer recognized choice options compared to unrecognized ones. Assuming that the memory strength of choice options is strongly correlated with both the choice criterion and recognition judgments, the RH is a reasonable strategy that approximates optimal decisions with a minimum of cognitive effort (Davis-Stober, Dana, & Budescu, 2010). However, theories of recognition memory are not generally compatible with this assumption. For example, some threshold models of recognition presume that recognition judgments can arise from two types of cognitive states: (1) certainty states in which judgments are almost perfectly correlated with memory strength and (2) uncertainty states in which recognition judgments reflect guessing rather than differences in memory strength. We report an experiment designed to test the prediction that the RH applies to certainty states only. Our results show that memory states rather than recognition judgments affect use of recognition information in binary decisions.

Keywords: recognition heuristic, fast and frugal decision heuristics, models of recognition memory, high threshold models, memory states.

1 Introduction

The recognition heuristic (RH) is a simple decision rule for binary choice problems (Gigerenzer & Goldstein, 1996; Goldstein & Gigerenzer, 1999, 2002). According to the RH, whenever people are required to decide between a recognized and an unrecognized choice option from a suitable decision domain, they will infer that the recognized option scores higher on the choice criterion and thus choose it. For example, assume that a person is going to invest a certain amount of money in one of two German stocks, say, "Kontron" or "Fraport". If the person recognizes the stock name "Fraport" but not the name "Kontron", the RH predicts that this person should choose the "Fraport" stock. In fact, the majority of people tend to follow this prediction, not only in stock market investments but in many other ecological decision domains as well (e.g., Marewski, Gaissmaier, Schooler, Goldstein, & Gigerenzer, 2010; Pachur, Bröder, & Marewski, 2008; Pohl, 2006). Moreover, RH-consistent choices have been shown to be very successful in many decision domains, although the evidence for success of the RH is somewhat mixed in case of stock market investments (Andersson & Rakow, 2007; Borges, Goldstein, Ortmann, & Gigerenzer, 1999; Boyd, 2001; Frings, Serwe, & Holling, 2003).

The success of the RH depends on the *correlation* between recognition judgments and the to-be-optimized criterion values. More precisely, (1) the higher the recognition validity (i.e., the proportion of paired comparisons in which the recognized option actually scores higher on the criterion than the unrecognized option), and (2) the more often recognition discriminates between two choice options, the higher the overall proportion of correct decisions based on the RH (Gigerenzer & Goldstein, 1996; Goldstein & Gigerenzer, 1999, 2002). In everyday experience, persons, objects, products, and goods scoring high on several choice criteria are frequently encountered and cognitively processed in various ways (e.g., triggered by communications, TV programs, or newspaper reports). Hence, advantageous choice options, compared to disadvantageous options, will produce stronger memory representations in most people, thereby boosting name recognition and increasing the recognition validity (e.g., Goldstein & Gigerenzer, 2002; Newell & Fernandez, 2006; Schooler & Hertwig, 2005). This is especially plausible for "natural" ecological decision domains that play a major role in everyday life. It is for these domains that the RH theory has been shown to be most successful, much more so than for artificial and abstract domains often studied in the laboratory (Pachur et al., 2008; Pohl, 2006).

Manuscript preparation was supported by a grant from the Deutsche Forschungsgemeinschaft (SFB 504, Project A12). We thank Janina Hoffmann, Simone Malejka, and Berenike Waubert de Puiseau for their help during manuscript preparation.

*Corresponding author: University of Mannheim. Address: Lehrstuhl für Psychologie III, Universität Mannheim, Schloss, D-68131 Mannheim, Germany. Email: erdfelder@psychologie.uni-mannheim.de.

†University of Mannheim

The RH theory proposed by Goldstein and Gigerenzer (1999, 2002) has stimulated much research (see Gigerenzer & Goldstein, 2011, and Marewski, Gaissmaier, & Gigerenzer, 2010, for recent reviews). Surprisingly, with a few noticeable exceptions (e.g., Marewski, Gaissmaier, Schooler et al., 2010; Pachur & Hertwig, 2006; Pleskac, 2007; Schooler & Hertwig, 2005), almost all articles published on the RH so far have ignored theories and models of recognition memory (see also Dougherty, Franko-Watkins, & Thomas, 2008; Hilbig, 2010b). Showing that the RH is an ecologically rational and well-adapted choice strategy obviously requires a formal theoretical link between (1) the memory strengths of choice option names — a latent variable which is affected by environmental frequency and previous processing — and (2) binary recognition judgments for choice option names — an empirical variable which is assumed to affect decision behavior. Theories of recognition memory provide this missing link. Ignoring these theories is tantamount to treating recognition judgments as determined solely by memory strengths, a view that is at odds with empirical facts. As is well known, recognition judgments are also affected by other influences, usually summarized under the label "response bias" (e.g., Macmillan & Creelman, 2005; Snodgrass & Corwin, 1988).

The present article aims at evaluating the implications of an influential and successful model of recognition memory, the Two-High-Threshold (2HT) model (Bredenkamp & Erdfelder, 1996; Snodgrass & Corwin, 1988), for the RH theory. In the next section, we begin by briefly reviewing recognition memory models that might close the gap between memory processes and recognition judgments in the RH framework. We then show that the 2HT model is most attractive because it can easily be combined with the RH in a more general theoretical framework — the *memory state heuristic* (MSH). The MSH model makes simple and straightforward predictions presuming that (1) the 2HT model of recognition memory is correct and (2) the use of recognition information is an ecologically rational decision strategy that is well adapted to environments frequently encountered in everyday life (Section 3). In the subsequent Section 4, we derive predictions for binary choice data from the MSH framework based on a process-model interpretation of the 2HT theory. Some of these predictions have already been corroborated in previous research but have remained unexplained in the RH framework so far. Other MSH predictions are new. In Sections 5 and 6, we report an experiment designed to test both types of predictions. Finally, implications for future research on recognition-based decision heuristics are outlined in the Discussion (Section 7).

2 Models of recognition memory

Two classes of stochastic models have dominated theories of recognition memory: signal-detection models and threshold models (Coombs, Dawes, & Tversky, 1970, chap. 6; Macmillan & Creelman, 2005; Snodgrass & Corwin, 1988). Both classes have been generalized to handle different types of one-dimensional and multidimensional memory judgments (e.g., Banks, 2000; Batchelder & Riefer, 1990; Bayen, Murnane, & Erdfelder, 1996; Bröder & Meiser, 2007; DeCarlo, 2003a, 2003b; Erdfelder, Cüpper, Auer, & Undorf, 2007; Hautus, Macmillan, & Rotello, 2008; Klauer & Kellen, 2010; Klauer & Wegener, 1998; Meiser & Bröder, 2002; Meiser, Sattler, & Weisser, 2008; Rotello, Macmillan, & Reeder, 2004). They have also been integrated in more general computational theories of memory and decision making (e.g., Brandt, 2007; Dougherty, 2001; Dougherty, Gettys, & Ogden, 1999; Juslin & Persson, 2002; Murdock, 1997; Shiffrin & Steyvers, 1997). However, for the present purposes it suffices to consider the simple case of yes-no recognition. In yes-no recognition tasks, each test probe is classified as either recognized (*yes* response) or not recognized (*no* response). We will first discuss signal-detection models and then turn to threshold models.

2.1 Signal-detection models

Signal-detection (SD) models assume that memory strength is a continuous variable following two normal distributions, one with mean d' and standard deviation σ for test items previously processed (old items) and the other with mean 0 and standard deviation 1 for lure items not processed previously (new items). The mean difference d' serves as a measure of increase in memory strength due to prior processing. A criterion value c defines the boundary between *yes* and *no* judgments on the memory strength dimension. If the memory strength exceeds this criterion, a person recognizes the test item as old (*yes* response); otherwise the test item is called new (*no* response). The location of the criterion, parameter c, provides a measure of response bias, with small values of c representing liberal response bias and large values of c indexing conservative biases (Macmillan & Creelman, 2005). In recognition memory studies using receiver-operating-characteristic (ROC) curves, it is usually found that $\sigma > 1$, supporting the so-called unequal-variance SD model (Macmillan & Creelman, 2005).

Schooler and Hertwig (2005, pp. 625–626) and Pleskac (2007) discussed implications of the SD model for the RH theory. Consistent with our view, Schooler and Hertwig (2005, p. 625) understand Goldstein and Gigerenzer's (2002) RH theory "… not so much as a model of recognition, but rather as a model of how the products of the recognition process could be used to make decisions" (see also Gigerenzer, Hoffrage, & Goldstein, 2008, p. 234). Based on the ACT-R framework (Anderson & Lebiere, 1998), they extended the empirical scope of the RH theory by modeling the underlying memory processes explicitly. The most general version of their model proposes two decision criteria, τ_1 and τ_2, on the memory strength dimension (called "activation" in ACT-R). Just like parameter c in the SD model, the ACT-R parameter τ_2 discriminates recognized and unrecognized objects. The additional parameter τ_1 defines the lower bound for objects that can be discriminated in terms of their retrieval latencies (which are monotonically decreasing in their activation values). The default version of their model assumes $\tau_1 = \tau_2$. In addition to the predictions of the RH theory, this model predicts that, if both objects are recognized, the choice option recognized more quickly is chosen — a decision rule known as the fluency rule of the fluency heuristic (FH). Moreover, by allowing for $\tau_1 < \tau_2$, their model could in principle also capture choices between two unrecognized objects, one of which is processed more fluently than the other and therefore preferred. Note, however, that the FH comes into play only when recognition does not discriminate between two choice objects. According to Schooler and Hertwig (2005, p. 623, Table 2), once one object is recognized and the other is not, the RH dictates that the recognized object should be chosen irrespective of recognition latency.[1] In other words, the FH theory put forward by Schooler and Hertwig (2005) is an extension of the RH theory (see also Hertwig, Herzog, Schooler, & Reimer, 2008). The predictions of the former include the core predictions of the latter.

Rather than exploring implications of SD models for actual choice behavior, Pleskac (2007) analyzed the implications of SD models with respect to the inferential accuracy of RH-consistent choice behavior. More precisely, Pleskac (2007) investigated the relationship between the recognition cue validity α and the ecological validity *A*, that is, the probability of option *x* exceeding option *y* on the choice criterion, given that *x* was experienced and *y* was not experienced previously. It is obvious that the correlation ϕ between the experience variable (experienced vs. not experienced) and the recognition judgment (*yes* vs. *no*) must approach $\phi = 1$ with increasing recognition memory accuracy. Hence, the larger d', the better should α approximate *A*. However, by assuming that people (1) conform to the RH perfectly and (2) follow a Bayesian-observer response strategy in the recognition test,[2] Pleskac (2007, p. 386, Figure 5) was able to derive the mathematical relation between *A* and α exactly, separately for different values of d'. More importantly, he could also show how d' influences the relation between the proportion of correct inferences based on the RH and the proportion of experienced objects in a certain decision domain. Specifically, the "less-is-more effect" — an inverted U-shaped relation between overall decision accuracy and the proportion of experienced objects — generally holds for large values of d' only. With decreasing memory sensitivity, the less-is-more effect diminishes and gives way to a "more-is-more effect", that is, the more objects were experienced the higher the inferential accuracy. In a nutshell, Pleskac's SD analysis showed that "... recognition ability plays a crucial role in the performance of the recognition heuristic" (Pleskac, 2007, pp. 389–390).

2.2 Threshold models

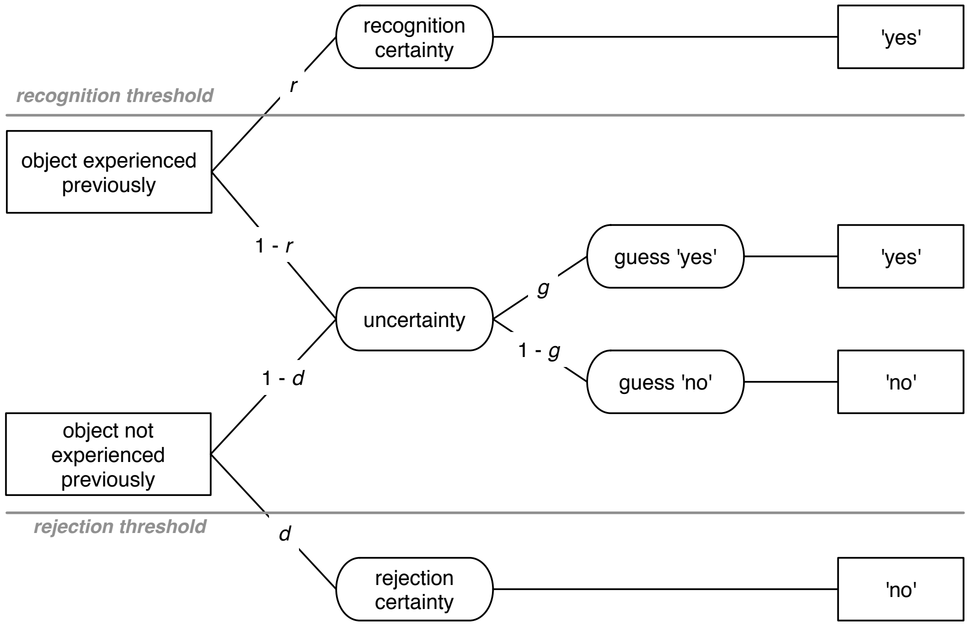
In contrast to signal-detection models, threshold models of recognition presume that recognition judgments depend on discrete memory states rather than continuous memory strength variables (Coombs et al., 1970, chap. 6). Of course, both conceptions are not necessarily incompatible, because discrete states can always be defined by placing categorization boundaries on continuous memory strength dimensions (Macmillan & Creelman, 2005). According to threshold models, however, possible differences in memory strength within memory states must not affect recognition judgments. In other words, only memory states matter for responding, not the strengths of underlying memory traces.

Some recognition memory theorists have argued that threshold models of recognition are inconsistent with the available empirical evidence because they imply linear ROC curves. As is well known, recognition memory ROC curves are usually found to be curvilinear (e.g., Wixted, 2007; Yonelinas & Parks, 2007). However, this argument ignores the important distinction between ROC curves based on confidence ratings and ROC curves based on yes-no recognition judgments for different experimental manipulations of response bias. Threshold models indeed predict linear ROCs for yes-no judgments but can easily account for curvilinear rating ROCs by using appropriate response models that map memory states on confidence ratings (Bröder & Schütz, 2009; Erdfelder & Buchner, 1998; Klauer & Kellen, 2010; Malmberg, 2002; Schütz & Bröder, in press; see also Batchelder

[1] More recently, however, alternative versions of the RH theory have been proposed that allow for the possibility that RH use is influenced by retrieval fluency (i.e., recognition latency) in recognition cases as well (Marewski, Gaissmaier, Schooler et al., 2010).

[2] Bayesian observers make use of a recognition criterion that is consistent with a posterior odds ratio of 1 for yes vs. no recognition judgments. In other words, they respond "yes" on the recognition test, if and only if the posterior probability of a test item being old exceeds the posterior probability of this test item being new.

Figure 1: The two-high threshold model for yes-no recognition tests (Snodgrass & Corwin, 1988). The two processing trees illustrate the cognitive processes leading to *yes* and *no* responses for old and new choice options, respectively. Ovals indicate latent cognitive states, and parameters attached to the branches denote (conditional) transition probabilities from left to right (r = probability of old objects exceeding the recognition threshold, d = probability of new objects falling below the rejection threshold, g = conditional probability of guessing *yes* in the uncertainty state).



& Riefer, 1999; Erdfelder et al., 2009). Because most published recognition memory ROC data are based on confidence ratings, the empirical evidence is not conclusive. The few relevant publications that use yes-no judgments in combination with experimental manipulations of response bias clearly rule out the so-called one-high threshold model (Swets, 1961) and the SD model assuming equal variances of old and new items. However, as shown convincingly by Bröder and Schütz (2009), the so-called two-high threshold model of recognition (Snodgrass & Corwin, 1988) can account for the available data at least as well as the unequal-variance SD model.

Surprisingly, the two-high threshold (2HT) model of recognition, although clearly one of the two best-established and best-supported formal recognition measurement models currently available (Bayen et al., 1996; Bredenkamp & Erdfelder, 1996; Bröder & Schütz, 2009; Erdfelder et al., 2007; Klauer & Kellen, 2010; Klauer & Wegener, 1998; Meiser & Bröder, 2002; Schütz & Bröder, in press; Snodgrass & Corwin, 1988), has been ignored in the RH literature so far. To the best of our knowledge, only Schooler and Hertwig (2005) made use of threshold models of recognition in their ACT-R account of the recognition and fluency heuristics (see also Marewski, Gaissmaier, Schooler et al., 2010). However, their model is based on a single threshold with respect to the yes-no recognition judgment. It also does not specify whether this threshold is "low" or "high" (i.e., whether it can or cannot be exceeded by new items that have not been experienced before; see Coombs et al., 1970, chap. 6). As a consequence, this model cannot be evaluated empirically based on ROC data. As stated explicitly by Schooler and Hertwig (2005, p. 617), their model "… does not produce ROC curves of any sort, simply because no mechanisms were specified to handle changes in response bias or to generate confidence ratings."

What can be gained by modeling response bias explicitly in threshold models of recognition? We briefly describe the 2HT model for yes-no recognition tests before we provide answers to this question.

The 2HT model belongs to the class of Multinomial Processing Tree (MPT) models (Batchelder & Riefer, 1999; Erdfelder et al., 2009) and can be illustrated in the form of a processing tree diagram (see Figure 1). The two processing trees in Figure 1 illustrate the possible sequences of cognitive processes for previously experienced objects (upper tree) and objects not experienced previously (lower tree). When processing an old item experienced previously, its memory strength may be so high that it exceeds the recognition threshold (probability r), producing a *yes* response (a "hit"). If it does not exceed the recognition threshold (probability $1 - r$), some kind of guessing (e.g., strategic guessing, informed guessing, random response choice, etc.) is required that may also lead to a correct *yes* response with probability g (a "hit" based on guessing) or to an incorrect *no* response with probability $1 - g$ (a "miss"). Thus, there are two possible process sequences resulting in correct *yes* responses for old items: (1) above-threshold recognition followed by a correct response (probability r) and (2) below-threshold uncertainty followed by guessing *yes* (probability $(1-r) \cdot g$). Because both events are disjoint, their probabilities sum up to the total probability of a hit:

$$p(\text{hit}) = r + (1 - r) \cdot g \tag{1}$$

The assumptions for new items are analogous to those for old items. When processing a new item, its memory strength may be so low that it falls below the rejection threshold (probability d)[3], thus immediately producing a *no* response (a "correct rejection"). If it does not reach the rejection threshold (probability $1 - d$), the new item is processed like an old item falling below the recognition threshold. That is, some kind of guessing is required that may lead to a wrong *yes* response (a "false alarm") with probability g or to a correct *no* response with probability $1 - g$ (a "correct rejection" based on guessing). Thus, there is only one possible process sequence resulting in a *yes* response for new items: an uncertainty state (probability $1-d$) followed by incorrectly guessing "yes" (probability g). Hence, the probability of a false alarm is:

$$p(\text{false alarm}) = (1 - d) \cdot g \tag{2}$$

The 2HT model equations (1) and (2) explain hit and false alarm rates in terms of three latent parameters: (1) the probability r of exceeding the recognition threshold, resulting in a *recognition certainty state*; (2) the probability d of falling below the rejection threshold, resulting in a *rejection certainty state*; and (3) the probability g of guessing *yes* in the *memory uncertainty state* which occurs whenever the memory strength falls in between the two thresholds. Note that both thresholds of the 2HT model are "high" in the sense that the recognition threshold can never be exceeded by new items and that the rejection threshold can never be reached by old items (Coombs et al., 1970, chap. 6).

3 The memory state heuristic

As outlined in the previous section, the 2HT model is a suitable measurement model of recognition memory. It nicely fits empirical ROC curves based on yes-no recognition judgments (e.g., Snodgrass & Corwin, 1988; Bröder & Schütz, 2009). Thus, there is good reason to assume that the 2HT model describes recognition memory states correctly. Importantly, fast and frugal decision heuristics like the RH and the FH are based on the principle of ecological rationality (Goldstein & Gigerenzer, 2002). Therefore, given the cognitive mechanisms of a typical decision maker, they should be well adapted to environments and contexts in which these decisions usually take place. Combining the 2HT model with the notion of ecological rationality, we should therefore try to find an answer to the following question: How would a well-adapted decision maker with a 2HT recognition memory use recognition information for decisions between choice options?

Let us assume that the decision domain is appropriate for the RH, that is, there is a strong correlation between the choice criterion and the memory strength of choice options. In such a domain, so-called "recognition cases" are most important for the RH. By definition, recognition cases are critical pairs (*x*, *y*) of choice options with *x* being recognized and *y* being unrecognized (Goldstein & Gigerenzer, 2002)[4]. According to the 2HT model, exactly four memory-state combinations may underlie such recognition cases: (*a*) recognition certainty for *x*, rejection certainty for *y*, (*b*) recognition certainty for *x*, uncertainty for *y*, (*c*) uncertainty for *x*, rejection certainty for *y*, and (*d*) uncertainty for both *x* and *y*. These four cases combine to three cases which need to be distinguished for our purposes: (1) both choice options are in certainty states (= case *a*), (2) both choice options are in the uncertainty state (= case *d*), and (3) one option is in a certainty state (either recognition or rejection certainty) and the other is in the uncertainty state (= cases *b* and *c*). Choice predictions based on the 2HT model are as follows:

1. *Both choice options are in certainty states*. The well adapted decision maker would most likely go with the recognition cue if *x* and *y* were in the recognition and rejection certainty states, respectively. Across these two states, recognition judgments are strongly correlated with memory strengths and thus very useful cues. Hence, consistent with the predictions of the RH, the recognized option *x* should be chosen.

2. *Both choice options are in the uncertainty state*. If both options were in the uncertainty state (with *x* and *y* receiving *yes* and *no* guesses, respectively), the well-adapted decision maker would most likely

[3] In many applications of the 2HT model, $r = d$ is assumed implicitly or explicitly (Snodgrass & Corwin, 1988). However, this assumption is not required in the 2HT framework although it often produces good fits to the data (e.g., Bayen et al., 1996; Klauer & Wegener, 1998; Meiser & Bröder, 2002). Importantly, however, given ROC data based on experimental manipulations of response bias, the assumption $r = d$ is not necessary to render the model identifiable or testable.

[4] Usually, the recognition data are obtained in separate yes-no recognition tasks performed independently of the decision task of choosing the option scoring higher on the criterion.

ignore the recognition cue, simply because recognition is determined solely by guessing in this state and thus uncorrelated to memory strength and the choice criterion. Hence, the probability of preferring *x* over *y* should be close to .50 in this case.

3. *Exactly one choice option is in a certainty state whereas the other is in the uncertainty state.* Following the same reasoning, the tendency to follow the recognition cue should be intermediate if one option is in a certainty state (either recognition certainty or rejection certainty) whereas the other is in the memory uncertainty state. This should hold irrespective of whether *x* is in the recognition certainty state and *y* is in the uncertainty state (i.e., a *no* guess) or, conversely, *x* is in the uncertainty state (i.e., a *yes* guess) and *y* is in the rejection certainty state. In both cases, recognition memory suggests that *x* scores somewhat higher than *y* on the choice criterion although the evidence is less compelling than in case (1). Hence, the probability of choosing *x* should be greater than .50 but clearly less than in case (1).

These predictions are simple and straightforward. Assuming that (a) the 2HT model is correct and that (b) humans are able to use memory state information to optimize choice behavior even if they are not consciously aware of these states (see Batchelder & Batchelder, 2008; Bayen, Nakamura, Dupuis, & Yang, 2000; and Meiser, Sattler, & von Hecker, 2007, for arguments and evidence pointing in this direction), our three predictions are implied whenever differences in memory states are strongly correlated with the choice criterion.

We call our model the memory state heuristic (MSH). Basically, the MSH can be seen as a three-state recognition heuristic which differs from the two-state, all-or-none RH in two important aspects: First, the MSH assumes a third state of *memory uncertainty* besides *recognition* and *rejection*. Second, in contrast to the RH, not all recognition cases are treated alike in the MSH. Rather than always choosing the recognized object, the MSH differentiates between possible underlying memory states. The MSH resembles the RH only if one object is in the recognition certainty state whereas the other is in the rejection state. For other possible memory state combinations, the MSH predicts less reliance on recognition (i.e., certainty-uncertainty pairs) or even ignorance of the recognition cue (i.e., uncertainty-uncertainty pairs).

Note that the MSH includes the RH of Goldstein and Gigerenzer (2002) as a special case: If the two threshold parameters r and d approach $r = d = 1$, then the uncertainty state vanishes. By implication, the MSH would then reduce to the RH, and the predictions of the RH and the MSH would become indistinguishable. We thus arrive at a result that reminds us of Pleskac's (2007) conclusion: In suitable domains, the higher the memory accuracy the more closely Goldstein and Gigerenzer's (2002) RH approximates well-adapted decision behavior.

4 Testing the memory state heuristic

We will report tests of predictions of the MSH as outlined in the previous section. Before doing so, we would like to point out that the MSH includes predictions not only for "recognition cases" but also for "guessing cases" and "knowledge cases" (i.e., cases in which neither option is recognized or both options are recognized, respectively; see Goldstein & Gigerenzer, 2002). The derivations of predictions for guessing and knowledge cases proceed analogous to those presented in the previous section. They require distinguishing between the memory state combinations that may underlie each of these cases (i.e., certainty/certainty, uncertainty/uncertainty, and certainty/uncertainty), resulting in three predictions for each of the two additional cases. Ideally, all nine predictions (three for recognition cases, three for knowledge cases, and three for guessing cases) should be incorporated into a single formal model of the MSH, analogous to the r-model recently suggested by Hilbig, Erdfelder, and Pohl (2010; see also Hilbig, 2010a; Hilbig & Richter, 2011; Hilbig, Scholl, & Pohl, 2010). However, the r-model, like the RH, refers to recognition judgments, not to memory states. Extending the r-model to recognition memory states (as required by the MSH) is clearly a nontrivial enterprise, since the experience variable (object experienced vs. not experienced) is not under the experimenter's control in typical applications. As recently shown by Bernstein, Rudd, Erdfelder, Godfrey, and Loftus (2009), it is possible in principle to fit memory models to recognition judgments even when the veracity of these judgments is unknown. However, although formal models should be the ultimate goals of psychological theorizing (e.g., Marewski & Olsson, 2009), it is certainly wise to start with simpler qualitative tests of the basic assumptions before making an effort of constructing complex stochastic models of the MSH including recognition, guessing, and knowledge cases.

Qualitative tests of the MSH can make use of the fact that MPT models like the 2HT model correspond to probabilistic serial processing models of cognition (Batchelder & Riefer, 1999; Hu, 2001). More precisely, each branch of the model's processing tree diagram can be seen as a possible temporal sequence of processing stages (see Figure 1 for the 2HT model). By implication, the total processing time per branch is the sum of the processing times for each of the cognitive stages along the branch. Hence, in the framework of the serial processing interpretation, MPT models entail implications concerning response time distributions in addition to response probabilities (see Hu, 2001).[5]

Following this reasoning, the 2HT model can be seen as a two-stage serial processing model analogous to the feature-comparison model of semantic memory judgments proposed by Smith, Shoben, and Rips (1974). In Stage 1, a recognition process evaluates the memory strength of choice options (processing time R). If the memory strength exceeds the recognition threshold, a fast *yes* response is given, and if it falls below the rejection threshold, a fast *no* response is given. A second processing stage is required only if the memory strength is intermediate (uncertainty state). In Stage 2, a guessing process generates responses in the state of memory uncertainty (processing time G). By implication, responses originating from the uncertainty state necessarily include two processing stages (overall processing time $R + G$), whereas certainty responses require a single processing stage only (processing time R). Hence, the distribution of response times originating from guessing must be stochastically larger than the response time distribution originating from memory certainty. A cumulative distribution function (cdf) $F(x)$ is called stochastically larger than a cdf $H(x)$ if and only if $F(x) \leq H(x)$ holds for all values of x. If X and Y are two real-valued random variables with values $x > 0$ and $y > 0$, respectively (e.g., x = processing time in the recognition stage, y = processing time in the guessing stage), then it follows that the cdf of their sum $z = x + y$, $F(z)$, is stochastically larger than the cdf of the first component x, $H(x)$ (see, e.g., Pfanzagl, 1991, p. 130). A direct implication is that the mean, the median, and all other percentiles of the response time distribution for judgments based on guessing exceed the corresponding parameters of the response time distribution based on memory certainty. In other words, the two-stage 2HT model predicts a strong correlation between memory states and recognition judgment latencies: The larger the recognition judgment latencies, the more likely it is that the judgment originates from guessing and the less likely it is that it originates from memory certainty.

Given the strong association between memory states and judgment latencies, three implications of the MSH for recognition cases are obvious:

1. The RH accordance rate (i.e., the probability of choosing the recognized object) should decrease with the recognition latency of the recognized object.

2. The probability of choosing the recognized object should decrease with the rejection latency of the unrecognized object.

3. The effects described in Predictions 1 and 2 should combine additively. By implication, the probability of choosing the recognized object should be highest when both judgment latencies are short (i.e., the *yes* and the *no* recognition judgment) and it should be lowest when both judgment latencies are long.

Prediction 1 was tested and confirmed in previous research, albeit in different theoretical frameworks (see Newell & Fernandez, 2006, Exp. 2.; Hertwig et al., 2008, Exp. 3; Marewski, Gaissmaier, Schooler et al., 2010). Results consistent with Prediction 1 have been argued to pose a challenge for a binary use of recognition in decision making as presumed by the RH (Newell & Fernandez, 2006, p. 342; Hertwig et al., 2008, p. 1199). As shown here, the negative correlation between the RH accordance rate and the recognition latency is not only unsurprising but, quite to the contrary, strictly implied by the MSH which assumes three recognition memory states rather than two.

To our knowledge, Predictions 2 and 3 are new and have not been tested before. It was one of the three purposes of the study reported in this paper to test these predictions and to replicate previous results concerning the first prediction using a different empirical paradigm. A second purpose was to test implications of the MSH with respect to the decision latencies between critical pairs of choice options, that is, the response latencies between presenting a pair of choice options and the act of making a choice. Decisions based on the RH have been argued to be particularly fast because they use a single piece of evidence only (“one-reason decision making”, Goldstein & Gigerenzer, 2002; see also Hertwig et al., 2008; Pachur & Hertwig, 2006). Hence, if the probability of using the RH increases with decreasing recognition and rejection latencies (Predictions 1 to 3) and if choices based on the RH are particularly fast, then the decision latencies should increase with the recognition and rejection latencies. We can thus derive three additional predictions for recognition cases analogous to Predictions 1 to 3:

4. The decision latency for critical pairs of choice options should increase with the recognition latency of the recognized object.

5. The decision latency for critical pairs of choice options should increase with the rejection latency of the unrecognized object.

[5]Note that the serial processing model is a reasonable albeit not mandatory interpretation of the 2HT model. Alternatively, the 2HT model can be seen as a pure measurement model which is silent about response latencies. However, we believe that the empirical scope of this model is enhanced substantially by adopting the serial processing interpretation outlined here. Similar interpretations of the 2HT model have previously been advocated by Atkinson and Juola (1974) and, more explicitly, by Bröder and Schütz (2009, pp. 599–600).

6. The effects described in Predictions 4 and 5 should combine additively, that is, the decision latencies for critical pairs of choice options should be shortest when both recognition judgment latencies are short (i.e., the *yes* and the *no* recognition judgment) and they should be longest when both judgment latencies are long.

Because response time analyses *between* individuals could easily produce spurious correlations as a consequence of individual differences in mental speed, Predictions 4 to 6 refer to response time differences *within* individuals.

A third and final purpose of our study was to assess effects of response bias in recognition judgments on choice behavior. Response bias manipulations are particularly interesting because they provide an additional tool to test between the original RH theory (Goldstein & Gigerenzer, 2002) and the MSH model. The original RH theory predicts that if response bias influences recognition judgments concerning choice option names then it should also affect choices between these options. This prediction derives from the fact that, according to the original RH theory, recognition judgments matter for decisions, not the underlying memory states. In contrast, the MSH predicts that only recognition memory states matter for decisions, not recognition judgments per se. Specifically, the probabilities of guessing *yes* or *no* in the uncertainty state must not affect choice probabilities. In the 2HT model framework, appropriate response bias manipulations should affect the guessing probability g selectively. Hence, they must not influence memory state probabilities (parameters r or d). In other words, the MSH implies the following prediction:

7. Response bias manipulations in the recognition test should affect recognition judgments but not binary decisions between choice options.

Importantly, in contrast to Predictions 1 to 6, Prediction 7 refers to all possible binary choices within a decision domain (or a representative sample of choices drawn from this domain). It does not necessarily hold when evaluated for recognition cases only, particularly in case of extreme response bias. The reason is that response bias may influence which memory state combinations enter into the set of recognition cases. For example, when response bias becomes extreme (i.e., either $g = 0$ or $g = 1$) then uncertainty-uncertainty pairs (i.e., type-2 pairs) vanish from the set of recognition cases. This problem does not occur when we consider an unselected, representative set of binary choices.

All seven predictions were tested in an experiment on recognition memory for German stock names and choices between pairs of these stocks.

5 Method

In order to test the abovementioned predictions, we manipulated the base rates of well-known and mostly unknown German stocks. One half of our participants was truthfully told that 60% of the stocks they would see in the study were well known whereas 40% of the stocks were quite unknown. Conversely, the other participants were truthfully told that 40% of the stock names would be well known whereas 60% would be unknown. Only real stock names were used in this study.

In addition to using different base rates, the order of the recognition phase and the decision phase was manipulated between participants.

5.1 Participants

One-hundred sixty-five persons participated in the experiment in exchange for monetary compensation and credit points. Participants received between three to five euros depending on their performance in the decision phase of the experiment (see Procedure section for details). Two of the participants had to be excluded from analysis due to technical failure during data collection, and two others because of a lack of *yes* judgments in the recognition phase.

One-hundred sixty-one participants remained in the set; 60.2% were female. Participants' ages ranged from 19 to 37 with a mean of 22.63 years. All participants were native German speakers except for one person whose language skills were nevertheless sufficient for the purpose of this study. Most of the participants were students (97.5%).

Because the study focused on stock investments, participants' stock knowledge was assessed as well. Therefore, participants had to evaluate their stock and stock-market knowledge. They could choose from a range of five options describing their knowledge as "very bad" (33.9%), "bad" (36.4%), "medium" (23.6%), "good" (4.8%) or "very good" (1.2%). In addition, participants were asked to state how often they would actively search for information on stocks and the stock market. Answers ranged from "never" (62.4%), "monthly" (23%), "weekly" (9.1%) to "daily" (5.5%).

5.2 Materials

Participants were presented with stock names that were either rated as "well-known" or "unknown" in a pilot study using 22 additional participants (not included in the sample described above). A total of 130 stocks listed on the German stock market homepage (http://deutsche-boerse.com) were assessed in the pilot study. Based on the results of the pilot study, sixty "well-known" and sixty "unknown" stocks were selected for the present study.

In the recognition phase of the experiment, participants were presented with 100 of these stock names. In the decision phase, these 100 names were arranged to form 50 stock pairs between which participants were asked to decide. Stock pairs consisted of either two "well-known" stocks, two "unknown" stocks, or of a "well-known" and an "unknown" stock. In the latter case the position of the "well-known" stock was randomly assigned to the left or right side of the computer screen. The assignment of stocks to position on the screen was predetermined. In addition, stock combinations fulfilled the criterion that one of the stocks was more successful than the other in the year preceding our study. All participants worked on the same stock pairs.

The base-rate information given to the participants did reflect the actual proportions of well-known and unknown stocks in the two phases. The material was counterbalanced across all experimental groups.

5.3 Design

Independent variables were the order of test phases (recognition phase first vs. decision phase first) and the base-rate manipulation (60% well-known stocks vs. 40% well-known stocks). Both factors were completely crossed and manipulated between subjects. In addition, two within-subject factors where defined for recognition cases, namely, recognition latency (short vs. long) and rejection latency (short vs. long). The terms "recognition latency" and "rejection latency" refer to the individual response times for *yes* and *no* judgments, respectively, in the recognition test. The medians of recognition and rejection latencies were calculated for each participant separately. Recognition and rejection latencies were classified as long when they exceeded the individuals' median recognition and rejection latencies, respectively. Otherwise they were classified as short. To summarize, our design was four-factorial with two dichotomous between-subjects factors (order of phases and base rates) and two dichotomous within-subject factors (recognition latency and rejection latency).

Dependent variables were relative frequencies and latencies of recognition judgments, the percentage of times participants chose stocks in accordance with the recognition heuristic (i.e., the accordance rate), the percentage of times participants chose the stock that was more successful in the year preceding the decision (i.e., the success rate), and the latencies of decisions in the decision phase.

5.4 Procedure

Participants completed the experiment in groups of up to 20 persons. They signed consent forms and were individually seated in front of a computer screen. All instructions used in the experiment were presented on the screen.

There were two phases in the experiment, the recognition phase and the decision phase. The order of these phases was randomized between participants. To make sure that the memory state of stock names does not change between the recognition phase and the decision phase, no break or distractor activity was inserted between the two phases. In the recognition phase, participants were presented with stock names that appeared sequentially in a random order in the middle of the screen along with the question "Do you know this stock?". For each stock name, participants had to press the red-labeled "D"-key for a "*yes*" response or the green-labeled "K"-key for a "*no*" response. Response speed was not emphasized in the instructions. Participants were not instructed to use specific fingers or hands for responding.

In the decision phase, participants were presented with pairs of stock names. The names were displayed in the middle of the screen along with the question "Which stock would you like to invest your money in?". Responses were provided by pressing the red-labeled "D"-key (indicating an investment in the stock shown on the left side) and the green-labeled "K"-key (indicating an investment in the stock shown on the right side of the display). Again, speed of responding was not emphasized.

During the decision phase, participants could collect additional credit points to increase the monetary compensation for their participation. Actual performances on the stock market were assessed for every stock. As a result participants' investments could be evaluated. Whenever they chose to invest in the stock (of a pair) that had done better from April 11, 2006 to April 11, 2007, they received additional credit points. Whenever participants chose to invest in the stock that had done worse in this time period, their score remained unchanged. Collecting more points resulted in achieving a higher monetary compensation.

All responses given in the experiment were self-paced. After completion of the experiment, participants were compensated for their participation and debriefed.

6 Results

For all statistical tests, $\alpha = .05$ was chosen as the criterion of statistical significance. Response time analyses always use medians of individual response times in the respective condition as dependent variables. In the following, we report means of these individual response time medians.

6.1 Overall response latencies

In the recognition phase, recognition and rejection judgments were faster when the decision phase preceded the recognition phase (997 ms) than when the recognition phase was worked on first (1235 ms). This difference was statistically significant ($t(159) = 6.078$, $p<.001$; Cohen's $d = 0.32$). Similarly, in the decision phase, decisions were faster when the recognition phase had been worked on before (2454 ms) than when the recognition phase followed the decision phase (3123 ms) ($t(159) = -4.892$, $p<.001$; Cohen's $d = 0.31$).

6.2 Overall recognition rates

If the 60% vs. 40% base rate manipulation of well-known stocks did in fact affect response bias, then, according to Prediction 7, it should also affect recognition rates for stock names. This was clearly the case. When the recognition phase preceded the decision phase, 33.64% of the 100 stock names were recognized in the 60% base-rate condition and only 26.60% were recognized in the 40% base-rate condition. Similarly, when the decision phase preceded the recognition phase, 35.90% were recognized in the 60% base-rate condition and only 30.60% were recognized in the 40% base-rate condition. A 2x2 ANOVA showed that only the base rate manipulation exhibited a significant main effect on the recognition rates ($F(1, 157) = 5.761$, $\eta^2 = .035$, $p = .018$) whereas effects of the order of phases and the interaction were not significant ($F(1, 157) = 1.459$, $\eta^2 = .009$, $p = .229$, and $F(1, 157) = .105$, $\eta^2 = .001$, $p = .747$, respectively). In other words, the recognition rate was larger in the liberal response bias condition, irrespective of the order of test phases. This result is consistent with Prediction 7 of the MSH.

6.3 Overall RH accordance rates

When considering the overall choice data for recognition cases, we can see that choices are consistent with the RH in the majority of cases. If the recognition phase preceded the decision phase, the RH accordance rates (probability of choosing the recognized object) were 75.1% and 79.3% for the 60% and the 40% base rate conditions, respectively. If the decision phase preceded the recognition phase, the RH accordance rates were 76.6% and 74.7% for the same two conditions. A 2x2 ANOVA showed that neither the order of phases nor the base rates nor their interaction significantly influenced the RH accordance rates (all Fs$(1, 157) \leq 1.687$, all $\eta^2 \leq .011$, $p \geq .196$). Given $N = 161$ and $\alpha = .05$, the power to detect medium deviations (**f** = .25; see Cohen, 1988) from H_0 is $1 - \beta = .884$ for all three F tests (Faul, Erdfelder, Buchner, & Lang, 2009), showing that insignificant results can safely be taken as evidence favoring H_0. Hence, we can conclude that the overall RH accordance rate is quite high and invariant against response bias manipulations.

6.4 Overall success rates

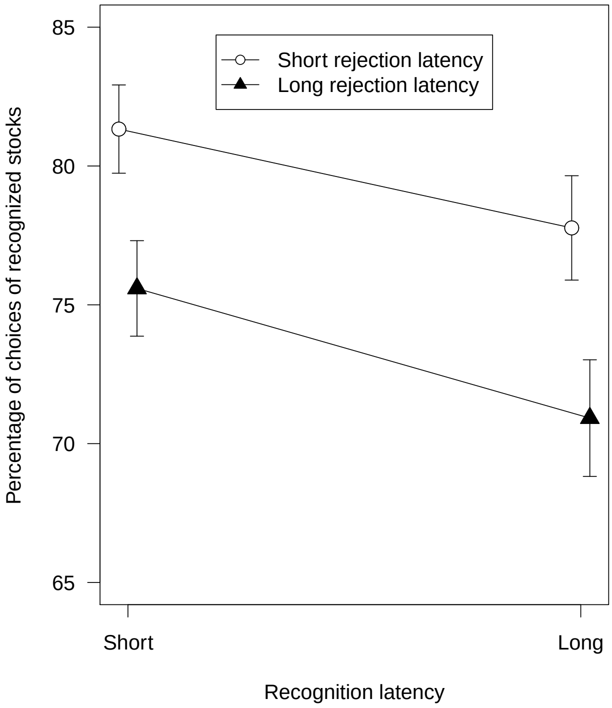
If the recognition phase preceded the decision phase, the percentages of choices of the more successful stock (irrespective of recognition status) were 48.7% and 49.1% for the 60% and the 40% base rate conditions, respectively. If the decision phase preceded the recognition phase, the corresponding success rates were 46.8% and 48.5% for the same two conditions. A 2x2 ANOVA showed that neither the order of phases nor the base rates nor their interaction significantly influenced the success rates (all Fs$(1, 157) \leq 1.351$, all $\eta^2 \leq .009$, $p \geq .247$). Hence, we can conclude that the overall success rate is invariant against response bias manipulations. As an aside, essentially the same results occur when we analyze a different indicator of choice behavior, namely, the percentage of choices of stocks presented on the left side of the computer screen, irrespective of recognition status or success of choices. Again, there was no indication of any differences between experimental groups in general and between response bias conditions in particular (all Fs$(1, 157) \leq 1.657$, all $\eta^2 \leq .010$, $p \geq .200$). Hence, choice behavior appears to be unaffected by response bias, consistent with Prediction 7.

Note that success rates of participants were at chance level in all experimental conditions. Essentially the same result is obtained when considering recognition cases only (i.e., the RH success rate) and when considering success in the year following the decision rather than success in the year preceding the decision. The low success rate may partly be due to the low level of stock expertise for the vast majority of our participants. In the stock domain, however, low success rates of the RH have been reported by other authors as well (Anderson & Rakow, 2007; Boyd, 2001; Frings et al., 2003).

6.5 RH accordance rates as a function of recognition and rejection latencies

Our first three predictions refer to the decrease in accordance rates for the recognition heuristic as a function of recognition and rejection latencies. As expected, a 2x2(x2x2) repeated measures ANOVA using order of phases and base rates as between-subjects factors as well as recognition and rejection latencies as within-subject

Figure 2: Mean RH accordance rates (and standard errors) in recognition cases as a function of recognition and rejection latencies.



factors showed that none of the between-subjects effects was significant (all F's(1, 154) $\leq$ 2.068, $\eta^2 \leq .013$, $p \geq .152$). In contrast, both the recognition latency ($F(1, 154) = 6.109$, $\eta^2 = .038$, $p = .015$) and the rejection latency ($F(1, 154) = 18.614$, $\eta^2 = .108$, $p < .001$) exhibited significant main effects, consistent with Predictions 1 and 2. Confirming Prediction 3, the interaction of recognition and rejection latency was not significant ($F(1, 154) = .126$, $\eta^2 = .001$, $p = .724$). Similarly, all interaction effects of the between-subjects and the within-subject factors were not significant (all F's(1, 154) $\leq$ 2.010, $\eta^2 \leq .013$, $p \geq .158$).

Figure 2 illustrates the mean RH accordance rates aggregated across the between-subjects factors as a function of recognition latency (short vs. long), separately for short and long rejection latencies. Error bars represent standard errors of the means. Obviously, the data pattern is in perfect agreement with Predictions 1 to 3. Looking at the main effects implied by the means shown in Figure 2, we see that the RH accordance rate decreases on average from 78.5% for short recognition latencies to 74.3% for long recognition latencies (aggregated across rejection latencies) and from 79.5% for short rejection latencies to 73.3% for long rejection latencies (aggregated across recognition latencies).

6.6 Decision latencies as a function of recognition and rejection latencies

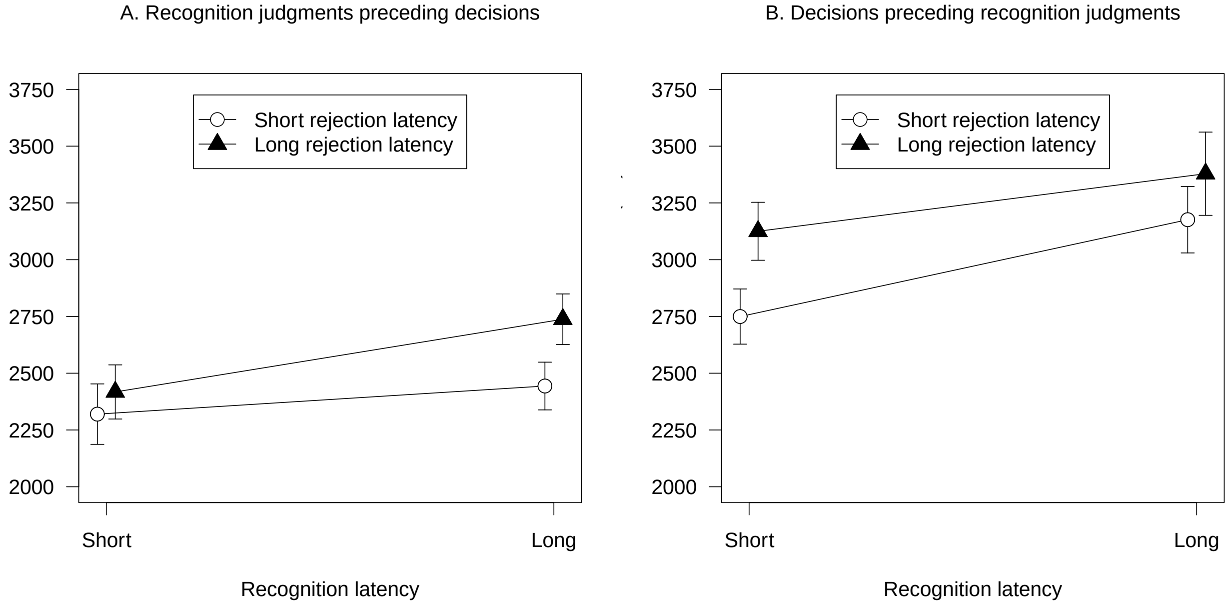
Predictions 4 to 6 refer to the increase in decision latencies as a function of recognition and rejection latencies. Again, a 2x2(x2x2) repeated measures ANOVA using order of phases and base rates as between-subjects factors as well as recognition and rejection latencies as two within-subject factors showed that, among the between-subjects effects, only the order of phases affected decision latencies significantly ($F(1, 154) = 18.86$, $\eta^2 = .109$, $p < .001$). None of the other between-subjects effects was significant (all F's(1, 154) $\leq$ 0.073, $\eta^2 \leq .001$, $p \geq .787$). In contrast, both the recognition latency ($F(1, 154) = 19.365$, $\eta^2 = .112$, $p < .001$) and the rejection latency ($F(1, 154) = 15.248$, $\eta^2 = .090$, $p < .001$) exhibited significant main effects consistent with Predictions 4 and 5. Confirming Prediction 6, the interaction of recognition and rejection latency was not significant ($F(1, 154) = .062$, $\eta^2 < .001$, $p = .803$). Similarly, all interaction effects of the between-subjects and the within-subject factors were not significant (all F's(1, 154) $\leq$ 2.788, $\eta^2 \leq .018$, $p \geq .097$).

Figures 3a and 3b illustrate the decision latencies (and standard errors) aggregated across the base-rate conditions as a function of recognition latency (short vs. long), separately for short vs. long rejection latencies and different phase orders. The data pattern corroborates Predictions 4 to 6. Estimating the main effects implied by the means shown in Figure 3a, we obtain average increases in decision latencies from 2930.16 ms for short recognition latencies to 3272.38 ms for long recognition latencies and from 2953.06 ms for short rejection latencies to 3249.47 ms for long rejection latencies. Analogously, estimating the main effects implied by Figure 3b, we observe average increases in decision latencies from 2344.72 ms for short recognition latencies to 2591.46 ms for long recognition latencies and from 2351.56 ms for short rejection latencies to 2584.61 ms for long rejection latencies.

7 Discussion

Research on the recognition heuristic (RH) has largely ignored theories of recognition memory. In line with other researchers who previously investigated the influence of memory processes on the RH (e.g., Pleskac, 2007; Schooler & Hertwig, 2005), we showed that modeling recognition memory processes explicitly has two major benefits for RH research. First, complications and anomalies in previous research results may turn out to be easily explainable if memory accuracy and response bias are taken into account. Second, several new predictions can be derived that have not been tested so far.

Figure 3: Mean decision latencies in recognition cases (and standard errors) as a function of recognition and rejection latencies. Figure 3a: Recognition judgments preceding decisions; Figure 3b: Decisions preceding recognition judgments.



We explored both routes using the memory state heuristic (MSH), a framework combining the well-supported two-high threshold (2HT) model of recognition (Bröder & Schütz, 2009; Snodgrass & Corwin, 1988) and the basic ideas of fast and frugal decision making using the RH (Gigerenzer & Goldstein, 1996; Goldstein & Gigerenzer, 1999, 2002). In contrast to the memory models discussed by Pleskac (2007) and Schooler and Hertwig (2005), the MSH is based on a threshold model of recognition that distinguishes between three memory states: recognition certainty, memory uncertainty, and rejection certainty. The basic idea is that these latent memory states determine choice behavior, not recognition judgments per se. This idea is reasonable because if decision makers always used recognition memory judgments rather than recognition memory states as cognitive input for their decisions, they would necessarily violate the principle of ecological rationality: In the memory uncertainty state, recognition judgments are uncorrelated with the choice criterion. It would thus not qualify as well-adapted choice behavior if a person always followed the recognition cue blindly, that is, even in the state of memory uncertainty.

Several complications and unexpected results observed in previous RH research can easily be explained in the framework of the MSH. The decrease in the RH accordance rate with increasing recognition latency (Newell & Fernandez, 2006; Hertwig et al., 2008), for example, has been unexplained in the RH framework so far. According to the MSH, this effect is expected because long recognition latencies typically arise from uncertainty states in which people dispense with RH use. Similarly, the fact that RH accordance rates are typically larger when RH-consistent decisions are correct than when they are incorrect (Hilbig & Pohl, 2008; see also Pachur & Hertwig, 2006) can easily be accommodated by the MSH theory without assuming use of further knowledge over and above recognition memory states. The explanation is as follows: First, given suitable decision domains, choices consistent with the predictions of the RH theory are most often correct when they originate from recognition and rejection certainty states, and they are least often correct when they originate from uncertainty states. Second, according to the MSH, people choose the recognized information in the former case but are indifferent in the latter case. Third, if one option is in the recognition uncertainty state and the other in one of the certainty states, both the probability of choosing the recognized option and the probability of being correct are intermediate. Hence, aggregated across recognition memory states, the MSH theory predicts a positive correlation between (a) the probability of choosing the recognized object and (b) the probability of a correct decision. This is in perfect agreement with Hilbig and Pohl (2008).

Another problematic result refers to the fact that choice options for which participants indicate knowledge in addition to mere recognition (so-called R+ recognition judgments) tend to be chosen more often than merely recognized options without additional knowledge (so-called mR judgments; see, e.g., Hilbig & Pohl, 2008; Hilbig, Pohl & Bröder, 2009; Marewski, Gaissmaier, Schooler et al., 2010; Newell & Fernandez, 2006, Exp. 1; Pohl, 2006). This result is difficult to reconcile with the original RH theory of Goldstein and Gigerenzer (2002) because knowledge in addition to mere recognition should be inconsequential once the recognition cue discriminates between the two choice options. One way to address this problem is to assume that RH use is moderated by retrieval fluency. Because retrieval fluency is typically less pronounced in mere recognition cases, "… it may often be harder to rely on recognition …, resulting in lower recognition heuristic accordance rates" (Marewski, Gaissmaier, Schooler et al., 2010, p. 295). Another explanation is suggested by the MSH. According to the MSH, the RH accordance rate should be higher when the recognized object is in the certainty state than when it is in the uncertainty state. Moreover, it is plausible to assume that recognized objects with R+ judgments have a higher likelihood of originating from the recognition certainty state than objects receiving mR judgments. Hence, the former should be chosen more often than the latter.

In a similar vein, the decision time results reported by Hilbig und Pohl (2009) are easily accommodated by the MSH. Hilbig and Pohl found that decision times between choice option pairs do not depend so much on whether these pairs conform to recognition or knowledge cases as defined by Goldstein and Gigerenzer (2002) but on their difference in terms of 3-alternative forced-choice recognition confidence ratings (i.e., U = unrecognized, mR = merely recognized, R+ = recognized with additional knowledge). For example, decision times for mR-R+ pairs turned out to be shorter on average than U-mR pairs although the latter are recognition cases and the former are knowledge cases according to Goldstein and Gigerenzer (2002). The MSH is consistent with this result. Depending on the response criteria of the participants, mR-R+ pairs might provide stronger evidence that the R+ option is in the recognition certainty state than U-mR pairs provide evidence for the mR option being in the recognition certainty state. Since only the memory state matters for responding according to the MSH, it makes sense that participants choose the R+ option in the mR-R+ pair more frequently and faster than the mR option in the U-mR pair.

Other effects that prompted extensions of the RH framework in the past are captured by the MSH theory in a parsimonious way. For example, Hertwig et al. (2008) convincingly showed that if two choice options are both recognized (a "knowledge case" according to Goldstein and Gigerenzer, 2002), then the option recognized more quickly tends to be chosen. The authors explain this by assuming a second decision heuristic in addition to the RH, namely, the fluency heuristic (FH). Note, however, that the same result can easily be explained by the MSH: According to the MSH, the choice option recognized more quickly should be chosen simply because it is more likely in the recognition certainty state. In other words, presuming two separate heuristics—the RH and the FH—is not necessary in the MSH framework. A single decision heuristic operating on recognition memory states rather than recognition memory judgments suffices to explain the results.

In addition to accounting for results that are at odds with the original RH and providing more parsimonious explanations than the FH theory of Schooler and Hertwig (2005), the MSH model predicts several new results that have not been tested before. Six of these predictions were tested and corroborated in the present study: RH accordance rates decrease with increasing rejection latencies, an effect that combines additively with the effect of recognition latencies (Prediction 2 and 3). Moreover, decision latencies increase and combine additively with the recognition and rejection latencies (Predictions 4 to 6). Last but not least, response bias manipulations affect recognition judgments but not binary choices (Prediction 7).

In addition to these predictions, several other testable hypotheses can be derived. For example, treatments that enhance memory accuracy by increasing the threshold parameters r and/or d should increase the RH accordance rate. The reason is that increases in these parameters minimize the probability of uncertainty states in which recognition information is not a useful cue.

Moreover, when both choice options are unrecognized ("guessing cases"; Goldstein & Gigerenzer, 2002), the choice option rejected with the longer rejection latency should be chosen more often in subsequent decisions. The rationale behind this prediction is that the option rejected quickly has the higher likelihood of being in the rejection state which is correlated with low ecological frequency. In contrast, the option rejected slowly is more likely in the uncertainty state — which associated with somewhat higher ecological frequency — and thus should be chosen. However, we expect this effect to be weaker than latency effects in recognition cases. The reason is that recognition cases include pairs of items both of which are in the most extreme (i.e., recognition vs. rejection) certainty states. According to the MSH theory, these pairs produce the highest accordance rates. In contrast, guessing cases — just like knowledge cases — cannot include pairs of items from the two extreme memory states. Stated differently, items both in guessing and in knowledge cases either belong to the same recognition memory state or to adjoining states. This tends to diminish latency effects compared to recognition cases, consistent with what Hertwig et al. (2008) observed.

We conclude that a modification of the original RH theory — the addition of a third state of memory uncertainty to the recognition and rejection certainty states considered by Goldstein and Gigerenzer (2002) — remedies most of the empirical problems of the RH theory that have been reported in the literature and predicts new results, several of which have been corroborated in our study. Importantly, given the MSH theory, there is no convincing evidence that would force us to assume additional heuristics like the fluency heuristic (Hertwig et al., 2008) or weighted fluency (Marewski, Gaissmaier, Schooler et al., 2010) to account for the results. The data appear to be consistent with a single three-state recognition heuristic, a straightforward extension of Goldstein and Gigerenzer's (2002) theory based on the two-high threshold model of recognition memory (Bröder & Schütz, 2009; Snodgrass & Corwin, 1988).

References

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

Andersson, P., & Rakow, T. (2007). Now you see it now you don't: The effectiveness of the recognition heuristic for selecting stocks. *Judgment and Decision Making, 2*, 29–39.

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology: I. Learning, memory, and thinking* (pp. 243–293). Oxford, GB: Freeman.

Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science, 11*, 267–273.

Batchelder, W. H. & Batchelder, E. (2008). Metacognitive guessing strategies in source monitoring. In J. Dunlosky & R. A. Bjork (Eds.), Handbook of metamemory and memory (pp. 211–244). Hillsdale: LEA books.

Batchelder, W.H., & Riefer, D.M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97*, 548–564.

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

Bayen, U. J., Murnane, K., & Erdfelder, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 197–215.

Bayen, U. J., Nakamura, G. V., Dupuis, S. E., & Yang, C.-L. (2000). The use of schematic knowledge about sources in source monitoring. *Memory & Cognition, 28,* 480–500.

Bernstein, D. M., Rudd, M. E., Erdfelder, E., Godfrey, R., & Loftus, E. F. (2009). The revelation effect for autobiographical memory: A mixture-model analysis. *Psychonomic Bulletin & Review, 16*, 463–468.

Borges, B., Goldstein, D. G., Ortmann, A., & Gigerenzer, G. (1999). Can ignorance beat the stock market? In G. Gigerenzer, P. M. Todd, & ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 59–72). London: Oxford University Press.

Boyd, M. (2001). On ignorance, intuition, and investing: A bear market test of the recognition heuristic. *Journal of Psychology and Financial Markets, 2*, 150–156.

Brandt, M. (2007). Bridging the gap between measurement models and theories of human memory. *Zeitschrift für Psychologie - Journal of Psychology, 215*, 72–85.

Bredenkamp, J., & Erdfelder, E. (1996). Methoden der Gedächtnispsychologie [Methods of the psychology of memory]. In D. Albert & K.-H. Stapf (Hrsg.), *Gedächtnis* (Enzyklopädie der Psychologie, Themenbereich C, Serie II, Band 4, S. 1–94). Göttingen: Hogrefe.

Bröder, A., & Meiser, T. (2007). Measuring source memory. *Zeitschrift für Psychologie - Journal of Psychology, 215*, 52–60.

Bröder, A., & Schütz, J. (2009). Recognition ROCs are curvilinear--or are they? On premature arguments against the two-high-threshold model of recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 587–606.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed.). Hillsdale: Erlbaum.

Coombs, C. H., Dawes, R. H., & Tversky, A. (1970). *Mathematical psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Davis-Stober, C. P., Dana, J., & Budescu, D. V. (2010). Why recognition is rational: Optimality results on single-variable decision rules. *Judgment and Decision Making*, *5*, 216–229.

DeCarlo, L. T. (2003a). An application of signal detection theory with finite mixture distributions to source discrimination. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 767–779.

DeCarlo, L. T. (2003b). Source monitoring and multivariate signal detection theory, with a model for selection. *Journal of Mathematical Psychology, 47,* 292–303.

Dougherty, M. R. P. (2001). Integration of the ecological and error models of overconfidence using a multiple-trace memory model. *Journal of Experimental Psychology: General, 130*, 579–599.

Dougherty, M. R. P., Franco-Watkins, A. M., & Thomas, R. (2008). Psychological plausibility of the theory of probabilistic mental models and the fast and frugal heuristics. *Psychological Review, 115*, 199–211.

Dougherty, M. R. P., Gettys, C. F., & Ogden, E. (1999). Minerva-DM: A memory process model for judgments of likelihood. *Psychological Review, 106*, 180–209.

Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models: A review of the literature. *Zeitschrift für Psychologie - Journal of Psychology*, *217*, 108–124.

Erdfelder, E., & Buchner, A. (1998). Process dissociation measurement models: Threshold theory or signal-detection theory? *Journal of Experimental Psychology: General, 127*, 83–97.

Erdfelder, E., Cüpper, L., Auer, T.-S., & Undorf, M. (2007). The four-states model of memory retrieval experiences. *Zeitschrift für Psychologie - Journal of Psychology, 215*, 61–71.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analysis using G*Power 3.1: Test for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160.

Frings, C., Serwe, S., & Holling, H. (2003). Anwendung der recognition heuristic auf den Aktienmarkt: Ignorance cannot beat the Nemax50 [The use of the recognition heuristic in the stock market: Ignorance cannot beat the Nemax50]. *Wirtschaftspsychologie, 5*, 31–38.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650–669.

Gigerenzer, G., & Goldstein, D. G. (2011). The recognition heuristic: A decade of research. *Judgment and Decision Making, 6*, 100–121.

Gigerenzer, G., Hoffrage, U., & Goldstein, D. G. (2008). Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty, Franco-Watkins, and Thomas (2008). *Psychological Review, 115*, 230–237.

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: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109*, 75–90.

Hautus, M. J., Macmillan, N. A., & Rotello, C. M. (2008). Toward a complete decision model of item and source recognition. *Psychonomic Bulletin & Review*, *15*, 889–905.

Hertwig, R., Herzog, S. M., Schooler, L. J., & Reimer, T. (2008). Fluency Heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 34*, 1191–1206.

Hilbig, B. E. (2010a). Precise models deserve precise measures: a methodological dissection. *Judgment and Decision Making, 5*, 272–284.

Hilbig, B. E. (2010b). Reconsidering “evidence” for fast-and-frugal heuristics. *Psychonomic Bulletin & Review*, *17*, 923–930.

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.

Hilbig, B. E., & Pohl, R. F. (2008). Recognizing users of the recognition heuristic. *Experimental Psychology, 55*, 394–401.

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, 34,*1296–1305.

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.

Hilbig, B. E., & Richter, T. (2011). Homo heuristicus outnumbered: Comment on Gigerenzer and Brighton (2009). *Topics in Cognitive Science, 3*, 187–196.

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.

Hu, X. G. (2001). Extending general processing tree models to analyze reaction time experiments. *Journal of Mathematical Psychology*, *45*, 603–634.

Juslin, P., & Persson, M. (2002). PROBabilities from EXemplars (PROBEX): A “lazy” algorithm for probabilistic inference from generic knowledge. *Cognitive Science*, *26*, 563–607.

Klauer, K. C., & Kellen, D. (2010). Toward a complete decision model of item and source recognition: A discrete-state approach. *Psychonomic Bulletin & Review*, *17*, 465–478.

Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the “Who Said What?” paradigm. *Journal of Personality and Social Psychology, 75*, 1155–1178.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2. ed.). New York: Cambridge University Press.

Malmberg, K. J. (2002). On the form of ROCs constructed from confidence ratings. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 380–387.

Marewski, J. N., Gaissmaier, W., & Gigerenzer, G. (2010). Good judgments do not require complex cognition. *Cognitive Processing*, *11*, 103–121.

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

Marewski, J. N., & Olsson, H. (2009). Beyond the null ritual: Formal modeling of psychological processes. *Zeitschrift für Psychologie - Journal of Psychology*, *217*, 49–60.

Meiser, T., & Bröder, A. (2002). Memory for multidimensional source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 116–137.

Meiser, T., Sattler, C., & von Hecker, U. (2007). Metacognitive inferences in source memory judgements: The role of perceived differences in item recognition. *Quarterly Journal of Experimental Psychology, 60,* 1015–1040.

Meiser, T., Sattler, C., & Weisser, K. (2008), Binding of multidimensional context information as a distinctive characteristic of Remember judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 32–49.

Murdock, B. B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review, 104*, 839–862.

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

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.

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.

Pfanzagl, J. (1991). Elementare Wahrscheinlichkeitsrechnung [Elementary probability theory], 2nd ed., Berlin, Germany: De Gruyter.

Pleskac, T. (2007). A signal detection analysis of the recognition heuristic. *Psychonomic Bulletin & Review, 14*, 379–391.

Pohl, R. F. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19*, 251–271.

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum-difference theory of remembering and knowing: A two-dimensional signal-detection model. *Psychological Review, 111*, 588–616.

Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review, 112*, 610–628.

Schütz, J., & Bröder, A. (in press). Signal detection and threshold models of source memory. *Experimental Psychology*.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM--retrieving effectively from memory. *Psychonomic Bulletin & Review, 4*, 145–166.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*, 214–241.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*, 34–50.

Swets, J. A. (1961). Is there a sensory threshold? *Science,* 134, 168–177.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114*, 152–176.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychological Bulletin, 133*, 800–832.