

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/220032645>

Multinomial processing tree models: A review of the literature

Article · January 2009

CITATIONS

106

READS

1,350

6 authors, including:



Benjamin E Hilbig

Universität Koblenz-Landau

102 PUBLICATIONS **2,110** CITATIONS

[SEE PROFILE](#)



Morten Moshagen

Ulm University

55 PUBLICATIONS **1,321** CITATIONS

[SEE PROFILE](#)



Lena Nadarevic

Universität Mannheim

7 PUBLICATIONS **130** CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Perception of websites [View project](#)

Multinomial Processing Tree Models

A Review of the Literature

Edgar Erdfelder, Tina-Sarah Auer, Benjamin E. Hilbig, André Aßfalg,
Morten Moshagen, and Lena Nadarevic

University of Mannheim, Germany

Abstract. Multinomial processing tree (MPT) models have become popular in cognitive psychology in the past two decades. In contrast to general-purpose data analysis techniques, such as log-linear models or other generalized linear models, MPT models are substantively motivated stochastic models for categorical data. They are best described as tools (a) for measuring the cognitive processes that underlie human behavior in various tasks and (b) for testing the psychological assumptions on which these models are based. The present article provides a review of MPT models and their applications in psychology, focusing on recent trends and developments in the past 10 years. Our review is nontechnical in nature and primarily aims at informing readers about the scope and utility of MPT models in different branches of cognitive psychology.

Keywords: multinomial modeling, measurement models, measurement of cognitive processes, multinomial processing tree models, software

In a now classical article, Riefer and Batchelder (1988) proposed a class of substantively motivated stochastic models for categorical behavioral data which was relatively well known in statistical genetics at the time (e.g., Elandt-Johnson, 1971), but had received little attention in psychological research up to the 1980s. These models are now known as multinomial processing tree (MPT) models. About 10 years later, Batchelder and Riefer (1999) already identified no less than 30 published MPT models in the psychological literature, most of which were applied to different agendas in cognitive research. The present article provides an update of Batchelder and Riefer's review and focuses on models and their applications published in the past 10 years. Our review includes 70 MPT models and model variants from more than 20 research areas.

In the first section, we will present a brief conceptual outline of MPT models using a simple example to illustrate the basics and main advantages of this approach. Technical details will be omitted almost entirely because they have been described elsewhere (e.g., Batchelder & Riefer, 1999; Hu & Batchelder, 1994). The second section summarizes MPT models and their applications in different branches of cognitive psychology, with a special focus on models for various memory paradigms. In the third section, psychological applications of MPT models outside the realm of cognitive psychology will be briefly summarized. The fourth section describes recent developments, generalizations, and innovations in the statistical methodology of MPT models that might be useful for those interested in applying such models. The fifth and final section of our review provides a sketch of computer programs that are currently available for statistical analyses in the MPT framework, along with a summary of the main advantages of each program.

What is MPT Modeling About?

MPT models address categorical data based on the assumption that the sample frequencies observed for a well-defined set of responses follow a multinomial distribution (Read & Cressie, 1988). Importantly, however, MPT models do not simply *assess* the probabilities underlying these sample frequencies. Rather, like most mathematical models of cognition (Ulrich, 2009), MPT models aim at *explaining* them in terms of latent processes that determine behavior. Moreover, they provide *measures* (i.e., probability estimates) of these cognitive processes and *statistical tests* of the psychological assumptions involved in a model.

To illustrate, let us consider the one-high threshold (1HT) model (Swets, 1961), probably the simplest nontrivial MPT model. Like any MPT model, the 1HT model is tailored to a specific empirical paradigm, in this case either a detection or a recognition paradigm. In the latter case, participants first learn a set of items and later receive a yes-no recognition test comprising old items (targets) and new items (lures). Participants are required to respond *yes* and *no* to targets and lures, respectively. Hence, two samples of discrete data are observed for each participant, namely (1) frequencies of correct *yes* and incorrect *no* responses to targets (called *hits* and *misses*, respectively) and (2) frequencies of incorrect *yes* and correct *no* responses to lures (called *false alarms* and *correct rejections*, respectively). A multinomial model for target items – in case of only two categories also called a binomial model – assumes that hits and misses arise from observations independently drawn from a population characterized by constant probabilities $p(\text{hit})$ and $p(\text{miss})$ that sum up to 1. Analogously, false alarms and correct rejections are assumed to follow a binomial model with complementary

probabilities $p(\text{false alarm})$ and $p(\text{correct rejection})$. If, as in our example, more than one multinomial (or binomial) sample is observed and the observations in different samples are independent, their joint distribution is called a *joint multinomial distribution*.

Using well-known statistical techniques such as maximum-likelihood (ML) parameter estimation, it is straightforward to estimate $p(\text{hit})$ and $p(\text{false alarm})$ (and, by implication, $p(\text{miss}) = 1 - p(\text{hit})$ and $p(\text{correct rejection}) = 1 - p(\text{false alarm})$) from the observed frequencies (Read & Cressie, 1988). However, these techniques would merely describe the results. Instead, to provide an *explanation* of underlying processes, the 1HT model assumes that recognition judgments can arise from two cognitive states, a *recognition certainty state* characterized by above-threshold memory strength or an *uncertainty state* associated with below-threshold memory strength. Let r denote the unknown probability that a target item is in the recognition certainty state. In this case, the item correctly receives a *yes* response. If, by contrast, the target item is in the uncertainty state (with probability $1 - r$), processes such as educated guessing will determine the response, resulting in a *yes* response with an unknown probability g or a *no* response with probability $1 - g$. Thus, there are exactly two routes into the *yes* response for targets: (1) an above-threshold recognition state followed by a correct response (with probability r) and (2) a below-threshold uncertainty state followed by a lucky guess (with probability $(1 - r) \cdot g$). Because both routes are disjoint, their probabilities sum up to the total probability of a correct *yes* response to target items:

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

How can false alarms arise for lures? The 1HT model assumes that the memory threshold is “high” to ensure that preexperimental familiarity or other influences cannot push lures above this threshold. By implication, lures are always assigned to the uncertainty state. In other words, lures are always subject to the same judgment processes as below-threshold targets. Hence,

$$p(\text{false alarm}) = g. \quad (2)$$

Equations 1 and 2 define the model equations¹ of the 1HT model. They explain the hit and false alarm rates in terms of two basic processes, viz., “above-threshold recognition” (r) and “guessing yes” (g). Given these equations and a set of sample data, the unknown probabilities r and g can be estimated using standard statistical estimation techniques (e.g., Read & Cressie, 1988). Because the 1HT model is a saturated model for a single set of targets and lures, ML estimates of r and g are easily derived by inserting relative frequencies of hits and false alarms into Equations 1 and 2 and solving for the parameters:

$$\begin{aligned} \hat{r} &= \frac{\hat{p}(\text{hit}) - \hat{p}(\text{false alarm})}{1 - \hat{p}(\text{false alarm})}, \\ \hat{g} &= \hat{p}(\text{false alarm}). \end{aligned} \quad (3)$$

If the assumptions underlying the 1HT model hold, these parameters yield a deeper understanding of recognition memory successes and failures than do hits and false alarms. For example, when analyzed in the 1HT model framework, the reduced hit rate typically found in depressives is due to a conservative response bias (i.e., small g) rather than an impaired recognition memory (i.e., small r) (cf. Miller & Lewis, 1977). Hence, effects that appear to be memory distortions at first sight may turn out to be judgment biases when analyzed in the framework of the 1HT model. This insight has important implications for both cognitive theory and psychological practice in various fields of applied psychology (e.g., psychotherapy).

Figure 1 illustrates the 1HT model in the comprehensible form of an MPT diagram. The parameters attached to the branches denote transition probabilities between cognitive states in the test situation and serve as measures of the process probabilities of interest. The label “processing tree” points to the core idea that each node in the tree represents a cognitive state or process. Thus, each tree branch stands for a possible sequence of cognitive processes that mediate between inputs (left side of the tree) and responses (right side of the tree). The probability of each possible sequence is the product of all process probabilities involved in a branch. In fact, all MPT models reviewed in this article can be illustrated in the form of such a binary processing tree diagram.² Each node is followed by (no more than) two branches with probabilities θ_s and $(1 - \theta_s)$, $s = 1, \dots, S$, each of which is an element of $[0, 1]$ (Hu & Batchelder, 1994).

MPT analyses of observed response frequencies typically involve four steps: (1) verification of model identifiability, (2) parameter estimation, (3) goodness-of-fit assessment, and (4) comparisons between different model parameters.

- (1) *Verification of identifiability.* A model is called identifiable if its equations define a one-to-one mapping of parameter vectors in the set of possible category probability vectors. Local identifiability (i.e., identifiability in the neighborhood of the true parameter values underlying a sample) is a precondition for the existence of unique parameter estimates given a set of sample frequencies. If local identifiability is violated, parameter constraints or extensions of the set of category probabilities may help to solve the problem.
- (2) *Parameter estimation.* Given an identifiable model, the S parameters of the model can be estimated (including

¹ Strictly speaking, these are the nonredundant model equations only. The remaining equations are easily derived because $p(\text{miss}) = 1 - p(\text{hit})$ and $p(\text{correct rejection}) = 1 - p(\text{false alarm})$.

² The general model equation of binary MPT models includes some models that cannot be represented as trees (Hu & Batchelder, 1994). However, these models are less relevant in practice and can be ignored in the present context.

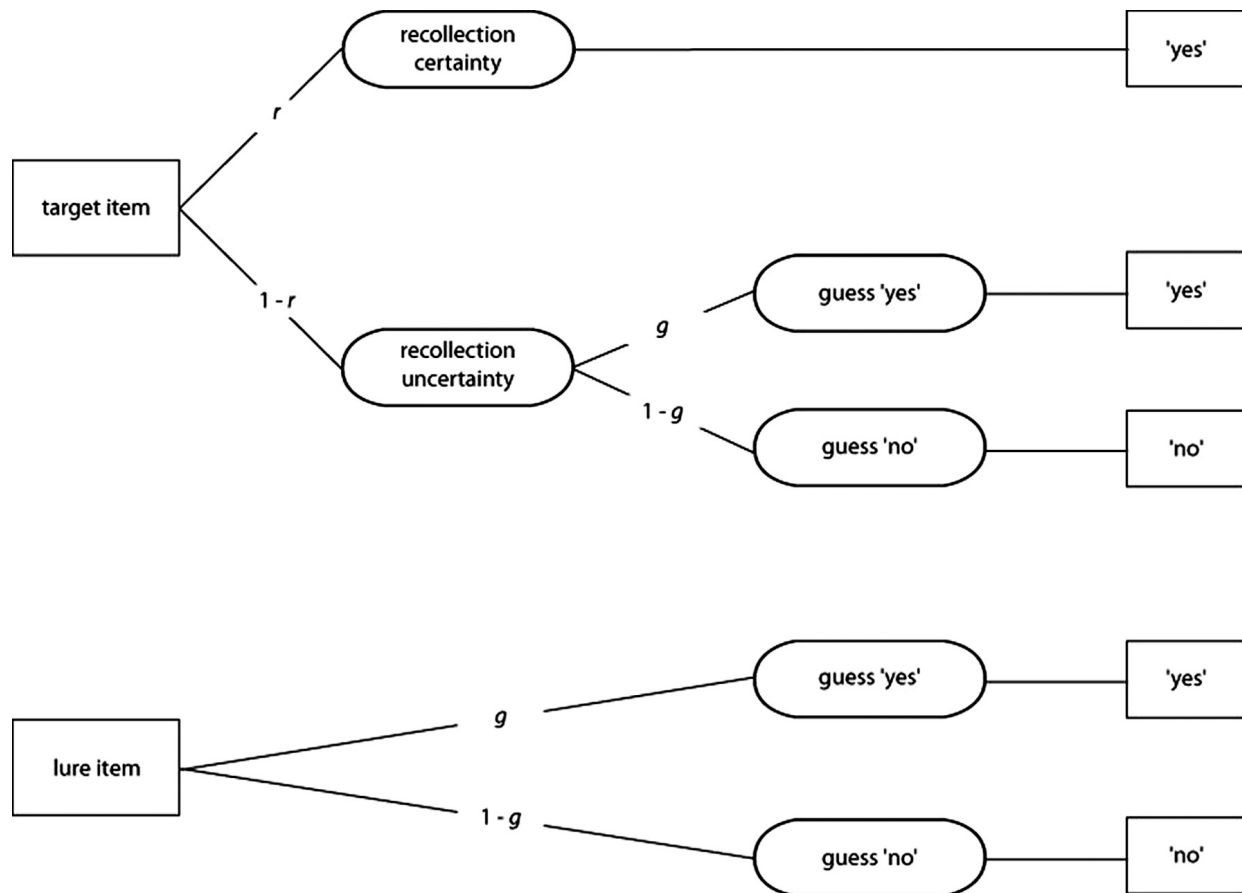


Figure 1. The 1HT model for yes-no recognition tests. The two processing trees illustrate the cognitive processes leading *yes* and *no* responses for targets and lures, respectively. Ovals indicate latent cognitive states and parameters attached to the branches denote (conditional) transition probabilities from left to right (r = probability of above-threshold memory strength and g = probability of guessing *yes* in the uncertainty state).

confidence intervals). This is typically achieved through minimizing a goodness-of-fit statistic, for example, by minimizing the log-likelihood ratio statistic $G^2(df)$ using the expectation-maximization algorithm suggested by Hu and Batchelder (1994). Minimum- G^2 estimates are equivalent to ML estimates and have several advantages (Read & Cressie, 1988).

- (3) *Goodness-of-fit test.* Third, the model's goodness-of-fit to the sample data is assessed, typically by comparing the goodness-of-fit statistic minimized in Step 2 against the relevant $\chi^2(df)$ reference distribution. If the observed fit statistic (e.g., G^2) falls below the $(1 - \alpha)$ percentage point of this distribution, the model is retained. Otherwise it is rejected at the α -level of significance. Alternatively, information-theoretic measures like the Akaike or Bayesian Information Criterion may be used to assess model fit (Read & Cressie, 1988). These statistics are especially helpful when comparing nonnested models.
- (4) *Parameter comparisons.* If an identifiable model fits the data, the final step typically concerns comparisons

between parameters of a model or tests of other parameter constraints (e.g., $g = .5$ in the 1HT model). Usually, such tests apply the G^2 difference statistic, ΔG^2 , which compares a superordinate model against a constrained nested submodel (Batchelder & Riefer, 1999). Such tests are of major importance for (1) validity evaluations of the model's parameters and (2) the interpretation of treatment differences or group differences in terms of underlying cognitive processes.

MPT Models in Cognitive Psychology

As noted above, MPT models were rarely used in psychology prior to the 1980s. Following the pioneering work of Riefer and Batchelder (1988), however, the situation changed dramatically, especially in the field of cognitive psychology. This section provides an overview of MPT models and their applications in cognitive psychology and social cognition research (see also Stahl, 2006). Because most of the published models refer to various memory paradigms, our focus will likewise be on memory models.

Recognition

Basic Models

Three types of MPT models for yes-no recognition and detection tests have been discussed in the literature: The 1HT model (Swets, 1961) as described in the previous section, the two-high threshold (2HT) model (Snodgrass & Corwin, 1988), and the low-threshold (LT) model (Krantz, 1969; Luce, 1963). Krantz (1969) also suggested more complex models with both high and low thresholds but these models have been less influential in subsequent years.

2HT models differ from 1HT models in assuming two thresholds that define three memory states: One threshold discriminating between recognition certainty (probability r) and uncertainty and a second threshold discriminating between distractor detection (probability d) and uncertainty. To make the model identifiable for a single pair of hits and false alarms, it is often assumed that $r = d$ (Snodgrass & Corwin, 1988). “Distractor detection” refers to metacognitive processes supporting the correct inference that a test item must be a lure (Strack & Bless, 1994). Hence, detected lures receive a *no* response, and guessing occurs for nondetected lures only. Both thresholds in the 2HT model are “high”, that is, lures always fall below the first threshold (as they do in the 1HT model) and targets always fall below the second high threshold (i.e., they are never detected as lures).

The LT model (Luce, 1963) is similar to the 1HT model in assuming only one threshold discriminating between two memory states. However, in contrast to the 1HT model, this threshold is “low”; thus, both targets and lures can cross this threshold, albeit with different probabilities.

For a single pair of hit and false alarm rates all three types of models are saturated and thus cannot be tested. However, by generalizing the models to *receiver operating curves* (ROCs), that is, to series of experimental conditions that differ in response bias (measured by g) only, they can be tested against each other. Both high threshold models imply simple linear relations between $p(\text{hit})$ and $p(\text{false alarm})$ for yes-no recognition tests in the ROC space, albeit with different slopes and intercepts. The empirical evidence available so far clearly rules out the 1HT model (Kinchla, 1994; Snodgrass & Corwin, 1988; Vorberg & Schmidt, 1975) but is roughly consistent with the 2HT model (Bröder & Schütz, 2009; Snodgrass & Corwin, 1988). The LT model can also account for typical data patterns (Luce, 1963) but implies a sharp corner in the ROC curve which is difficult to evaluate empirically. Thus, for the time being, the 2HT model appears to be the best two parameter MPT model for yes-no recognition memory tests.

Extensions

MPT models for recognition paradigms have been extended in various ways. For example, Erdfelder and Buchner (1998a), Malmberg (2002), and Bröder and Schütz (2009)

generalized threshold models to Likert-type confidence-rating scales. To discriminate between different types of recognition, Stenberg (2006) generalized the 1HT model to complex designs providing for measures of recognition based on perceptual form and recognition based on meaning.

Other popular extensions of recognition paradigms address the remember-know (RK) and remember-know-guess (RKG) procedures: Following a *yes* judgment in a yes-no recognition test, these procedures require participants to qualify their judgment as “remembered” (R), “known to be old without recollection” (K), or “guessed” (in the RKG procedure only). Xu and Bellezza (2001) evaluated dual-process extensions of the 1HT model to RK and RKG paradigms that provide separate measures of recollection and familiarity. However, their models failed to fit the data. Erdfelder, Cüpper, Auer, and Undorf (2007) showed that this problem can be solved by extending the 2HT model rather than the 1HT model to RK or RKG paradigms. They proposed a model of memory retrieval experiences based on four memory states: *recollection*, *familiarity*, *uncertainty*, and *rejection* (i.e., distractor detection).

Applications

Applications of threshold models, primarily the 2HT model, in recognition memory research have been so numerous that a comprehensive review is beyond the scope of this article. Many memory researchers routinely use the “corrected hit rate” $p_r = p(\text{hit}) - p(\text{false alarm})$ and the response bias parameter b_r . These parameters correspond to the recognition and guessing parameters of the 2HT model.

One important application concerns the problem whether MPT models can account for the typical ROC curve found in recognition experiments using confidence-rating scales. This ROC curve is curvilinear and becomes roughly linear only after z -transforming the hit and the false alarm rates. Many researches have taken this as evidence against MPT recognition models and as a confirmation of the signal-detection model with unequal variances (SD model; Green & Swets, 1966). As Erdfelder and Buchner (1998a), Malmberg (2002), and Bröder and Schütz (2009) have shown, this conclusion is premature. Extensions of 2HT models to confidence-rating scales can easily account for the curvilinear nature of the ROC curve typically found in these experiments. It is only for yes-no recognition tests that MPT threshold models imply linear ROC functions. The empirical evidence available so far (Bröder & Schütz, 2009; Snodgrass & Corwin, 1988) is consistent with these predictions. The 2HT model fits yes-no recognition ROC curves approximately as well as the unequal-variance signal-detection model does. Thus, the issue whether memory strength operates as a continuous or a discrete variable has not been settled so far.

Source Monitoring

Source monitoring tests can be understood as simple extensions of recognition tests. The test items stem from

two (or more) different sources, for example, from different speakers who previously presented the items in the study phase. In the recognition phase, participants are not only asked to judge whether an item is old or new but also to which source it belongs (for a review of source memory measurement methods, see Bröder & Meiser, 2007).

Basic Models

Batchelder and Riefer (1990) proposed the first MPT source monitoring model. Like the 1HT recognition model, Batchelder and Riefer's model assumes a single high threshold discriminating between recognition certainty and uncertainty. In addition to item recognition parameters it includes parameters representing source memory and various forms of guessing. Batchelder, Riefer, and Hu (1994) and Bayen, Murnane, and Erdfelder (1996) suggested LT and 2HT versions of the same basic model. A comparative empirical evaluation of all three model variants showed that only the 2HT version reflected all manipulations in a psychologically plausible way (Bayen et al., 1996). It is probably for this reason that most recent model extensions are based on the 2HT source monitoring model (see, however, Menon & Woodward, 2007).

Extensions

Riefer, Hu, and Batchelder (1994) proposed a 1HT model extension for source monitoring tasks with three or more different sources. This model requires less restrictive assumptions concerning guessing processes and is easily adapted to LT and 2HT frameworks. Another extension was developed to account for partial source memory (Dodson, Holland, & Shimamura, 1998). Partial source memory is available if participants can, for example, remember the sex of a speaker, but not which of two same-sex speakers presented an item. Klauer and Wegener (1998) independently developed a model for the "Who said what" paradigm in social categorization research. Their model is very similar to Dodson's model, with the main difference that Klauer and Wegener's model hinges on the 2HT assumption whereas Dodson and coworkers proposed a 1HT model variant.

Likewise, based on the 2HT assumption, Meiser and Bröder (2002) presented a model for crossed source dimensions, that is, two sources (e.g., color and position) with two attributes (e.g., blue, red; left, right). Participants are asked to judge both source attributes for each item. To account for different theoretical assumptions underlying this model, Meiser (2005) developed a model hierarchy based on the model for crossed sources. With a similar aim, Meiser and Bröder (2002) combined their model with RK judgments. This enabled them to test predictions about the stochastic dependence of different aspects of the encoding episode. Leaning on this approach, Dodson (2007) extended the simple 2HT source memory model to RK judgments. The most recent extension of the 2HT model was developed to

account for different guessing biases depending on the class of distractor items (Vogt & Bröder, 2007).

Applications

2HT source monitoring models have been used to address various research problems, for example, illusory correlations (Klauer & Meiser, 2000), source memory for faces (Bell & Buchner, in press; Buchner, Bell, Mehl, & Musch, 2009), and age differences in irrelevant speech effects (Bell, Buchner, & Mund, 2008). In clinical psychology, source memory models have been applied to study schizophrenia (Aleman, Bökler, Hijman, de Haan, & Kahn, 2003; Keefe, Arnold, Bayen, & Harvey, 1999; Keefe, Arnold, Bayen, McEvoy, & Wilson, 2002; Woodward, Menon, Hu, & Keefe, 2006), closed-head injury (Schmitter-Edgecombe, Marks, Wright, & Ventura, 2004), depression in college students (von Hecker & Meiser, 2005), and dementia in elderly people (Simons et al., 2002; for a review of clinical applications, see also Batchelder & Riefer, 2007). The specific and partial source memory in healthy older adults (Simons, Dodson, Bell, & Schacter, 2004) was studied using the model of Dodson et al. (1998).

Source memory models were also applied to further investigate the influences of schematic knowledge as well as stereotype information on item and source memory judgments (Bayen, Nakamura, Dupuis, & Yang, 2000; Dodson, Darragh, & Williams, 2008; Ehrenberg & Klauer, 2005; Erdfelder & Bredenkamp, 1998; Spaniol & Bayen, 2002). Dodson (2007) and Dodson et al. (2008) investigated illusory source recollections and the role of the retrieval process with Bayen's 2HT model. Klauer and Wegener's (1998) "Who said what?" source monitoring model has attracted a lot of attention especially in social categorization research (Gawronski, Ehrenberg, Banse, Zukova, & Klauer, 2003; Klauer & Ehrenberg, 2005; Klauer, Ehrenberg, & Wegener, 2003; Klauer, Wegener, & Ehrenberg, 2002; Wegener & Klauer, 2004, 2005; Wegener et al., 2008).

Meiser and Hewstone (2004, 2006) applied Meiser and Bröder's (2002) model to study stereotype formation and the role of illusory and spurious correlations for this process. The model was also used to study differences in the generation effect between reality monitoring and external source monitoring (Riefer, Chien, & Reimer, 2007). Meiser, Sattler, and von Hecker (2007) showed that metacognitive (subjective) knowledge about the recognizability of different sources can influence guessing parameters in multidimensional source monitoring models. Bröder, Nothen, Schütz, and Bay (2007) applied Vogt and Bröder's (2007) extended model to assess whether hidden covariations between sources and various stimulus attributes lead to specific guessing biases.

Moreover, the Meiser and Bröder (2002) model with RK judgments was used to unveil different binding processes in RK judgments (Meiser, Sattler, & Weißer, 2008). Specifically, stochastic dependence in the binding process for different context attributes was only found in remember judgments. The same model was applied to investigate source memory for perceptual details and the conditions

under which these form a basis for remember judgments (Meiser & Sattler, 2007).

Process Dissociation

The process dissociation procedure (PDP) was proposed by Jacoby (1991) to provide process-pure measures of controlled versus automatic retrieval processes (recollection vs. familiarity). Participants are required to provide judgments under two experimental conditions typically manipulated within subjects: In the inclusion condition, recollection and familiarity act in concert whereas in the exclusion test condition they act in opposition. Different variants of the PDP have been proposed, for example, for recognition tests, word-stem completion tests, fame judgment tasks, and cued-recall tasks.

Basic Models

The original PDP measurement model was developed by Jacoby (1991). It provides measures of controlled and automatic retrieval processes but cannot account for effects of response bias. Buchner, Erdfelder, and Vaterrodt-Plünnecke (1995) proposed an extended measurement model (EMM) that includes parameters representing guessing processes in the inclusion and exclusion conditions along with measures of controlled and automatic processes. The EMM assumes single high threshold for target items only. 2HT generalizations have been suggested that include an additional distractor detection parameter (Erdfelder & Buchner, 1998a; Meiser, 2000; Sherman, Groom, Ehrenberg, & Klauer, 2003; Vaterrodt-Plünnecke, Krüger, Gerdes, & Bredenkamp, 1996).

Extensions

Several extensions of the original PDP and the associated measurement models have been proposed. Erdfelder and Buchner (1998a) showed how to generalize the EMM to confidence-rating scales. This model version requires additional parameters but can account for rating-based inclusion and exclusion ROC curves just as well as the dual-process signal-detection model (e.g., Yonelinas & Jacoby, 1996). Jacoby (1998) proposed two different MPT models for the PDP based on the word-stem completion paradigm: the EMM (which he calls the “direct-retrieval model”) and a new “generate-recognize model”. Depending on the instructions, the latter model may provide a better fit to PDP word-stem completion data. Still other MPT models for PDP procedures have been proposed by McBride and Doshier (1999) and McBride, Doshier, and Gage (2001). Adaptations of several of these models to recall-then-recognition paradigms have been developed and tested by Bellezza (2003).

Vaterrodt-Plünnecke, Krüger, and Bredenkamp (2002) suggested an extension of the PDP to assess the independence assumption concerning controlled and automatic processes (cf. Erdfelder & Buchner, 2003; Hillstrom & Logan, 1997;

Rouder, Lu, Morey, Sun, & Speckman, 2008). The evidence available so far casts doubt on the independence assumption. Consequently, it seems prudent to apply measurement models like the EMM that do not require such an assumption (Buchner & Erdfelder, 1996; Stolz & Merikle, 2000).

Applications

PDP measurement models have been used repeatedly to assess controlled and automatic memory processes in various contexts (e.g., Bröder & Bredenkamp, 1996; Caldwell & Masson, 2001; Cüpper & Erdfelder, 2004; McBride & Shouder, 2003; Ott, Curio, & Scholz, 2000) as well as explicit and implicit sequence knowledge in implicit serial learning paradigms (Buchner, Steffens, Erdfelder, & Rothkegel, 1997; Buchner, Steffens, & Rothkegel, 1998). One methodologically important application concerns the relation between the cognitive processes involved in PDP paradigms and in source monitoring paradigms. It has been shown that PDP data can often be predicted from source monitoring data and vice versa (Buchner, Erdfelder, Steffens, & Martensen, 1997; Steffens, Buchner, Martensen, & Erdfelder, 2000). Hence, the cognitive processes underlying both paradigms appear to be the same. Because the source monitoring paradigm is more efficient and easier to apply, most PDP paradigms could be replaced by source monitoring paradigms. MPT measures of controlled and automatic memory processes, although not equivalent to measures of source and item memory, can be derived from source monitoring data (Buchner et al., 1997).

Conjoint Recognition

The conjoint recognition procedure was proposed by Brainerd, Reyna, and Mojardin (1999) as an alternative to the PDP. It is based on a recognition test consisting of three types of probes: (A) old items from the study list, (B) related lure items that share a target's gist, and (C) unrelated lure items. Participants receive one of the three instructions: (1) accept only items of Type A, (2) accept only items of Type B, or (3) accept items of Types (A) and (B).

Basic Model

Brainerd et al. (1999) developed and evaluated an MPT model for the conjoint recognition paradigm. Rather than addressing controlled and automatic retrieval processes as the PDP models do, their model provides core parameters representing the probabilities of retrieving *verbatim traces*, *gist traces*, and *recollection rejection* as defined in the authors' fuzzy trace theory.

Extensions

The model of Brainerd et al. (1999) has rarely been applied. Stahl and Klauer (2008) attributed this to the complexity of

the empirical paradigm. Consequently, they developed and evaluated a more efficient conjoint recognition paradigm and an associated measurement model.

Applications

Stahl and Klauer (2009) recently applied their model to study phantom recollection processes.

Storage-Retrieval Paradigms

A classical problem addressed by MPT models concerns the separation of storage and retrieval processes. In fact, some of the first MPT models referred to this topic (cf. Batchelder & Riefer, 1980; Chechile & Meyer, 1976). Extensions (Bröder, in press; Chechile, 2004) as well as applications of these early models are still of interest (Bröder, Herwig, Teipel, & Fast, 2008; Chechile, in press a).

Basic Models

Three basic models have been developed to disentangle storage and retrieval processes: the pair-clustering model by Batchelder and Riefer (1980, 1986), the trace-susceptibility model by Chechile (1987), and the model by Riefer and Rouder (1992). The pair-clustering model can be applied to free recall data of word lists that contain semantically clusterable word pairs as well as single words. The parameters represent probabilities of (1) storing two clusterable words as a cluster, (2) retrieving stored clusters, and (3) storing and retrieving nonclustered words. The trace-susceptibility model (Chechile, 1987) is based on a paradigm in which free recall and recognition trials are randomly intermixed. The use of a three-point confidence-rating scale in recognition trials allows for the estimation of storage, retrieval, and several guessing parameters. The model rests on the assumption that storage alone is sufficient for successful recognition, whereas both storage and retrieval are necessary to recall a target word. This assumption also underlies the third basic model by Riefer and Rouder (1992), tailored to a paradigm which includes free recall followed by cued-recall trials for word pairs. Comparable to the pair-clustering model, this model also estimates the storage of paired associates, the retrieval of associations, and hybrid parameters.

Extensions

Storage-retrieval models have been extended in various ways. For example, Chechile (2004) proposed a variant of the Chechile-Meyer model to overcome critical assumptions underlying the foil-recognition tree in the original model. Bender, Wallsten, and Ornstein (1996) suggested a storage-retrieval model for sequences of free recall tasks and applied it in a developmental context. Rouder

and Batchelder (1998) developed a family of nine nested models for the free-then-cued-recall paradigm also used by Riefer and Rouder (1992). Steffens, Jelenec, Mecklenbräuker, and Thompson (2006) and Steffens, Jelenec, and Mecklenbräuker (2009) recently proposed an extension of storage-retrieval models to combinations of free recall, cued-recall, and recognition tests to measure item-specific processing, relational processing, and retrieval in free recall. They applied their model to assess the cognitive mechanisms underlying the enactment effect in free recall. Finally, Bröder (in press) developed a model based on the pair-clustering model to disentangle storage and retrieval of word pairs as well as source memory for each item.

Applications

Beyond the original purpose of separating encoding and retrieval contributions to the spacing effect in free recall, the pair-clustering model was also used in developmental and clinical settings. Riefer and Batchelder (1991) revealed that older people have problems in retrieving word clusters. With higher statistical power, Erdfelder and Bayen (1991) found an additional storage deficit. A recent application of this model (Bröder et al., 2008) included an extended multi-trial version (Knapp & Batchelder, 2004) and compared healthy older adults and seniors with mild cognitive impairment. This analysis revealed globally decelerated learning and a pronounced retrieval deficit in the clinical population.

The pair-clustering model was also successfully applied in other clinical samples, for example, Alzheimer and vascular demented people (Batchelder, Chosak-Reiter, Shankle, & Dick, 1997), schizophrenics and alcoholics (Riefer, Knapp, Batchelder, Bamber, & Manifold, 2002), and stroke patients (Golz & Erdfelder, 2004). Batchelder and Riefer (2007) provide an according overview.

The models developed by Chechile and coworkers (cf. Chechile, 2004) were also used to investigate clinical groups (Chechile, in press a) and memory performance in different age groups (Chechile, 2007).

Models applicable to free-then-cued-recall data revealed that the bizarreness effect is due to a better retrieval of bizarre words (Riefer & Rouder, 1992), whereas common words are better stored in memory (Riefer & LaMay, 1998). The “recall of non recognizable words” phenomenon (Tulving & Thomson, 1973) was also analyzed in more detail with a multinomial model (Riefer & Batchelder, 1995). This analysis uncovered that violations of the Tulving-Wiseman function (Tulving & Wiseman, 1975) occur if weak storage is coupled with strong retrieval.

Immediate Serial Recall

The immediate serial recall task is probably the most popular paradigm in short-term and working memory research. After learning a sequence of items, participants are required to recall them immediately in exactly the same

sequence in which they were presented. Recalled items are scored as correct only if they are recalled in the correct position.

Basic Model

Schweickert (1993) proposed a two-parameter MPT model to assess the probability of intact item representations in short-term memory and the probability of successful reintegration, given a degraded memory trace. Both processes result in a correct serial recall of an item. Errors can occur only if both processes fail. The basic model has two parameters and only one nonredundant model equation. Hence, it cannot be identifiable. However, by generalizing the model to experimental designs in which parameters are selectively influenced, it is possible to achieve more independent model equations than parameters to estimate. Interestingly, the model is still not identifiable. However, the parameter estimates are unique up to certain types of linear transformations (for details, see Buchner & Erdfelder, 2005, Appendix); thus, at least the rank order of the parameters across factor levels can be interpreted which is sufficient for most applications.

Applications

The model has been applied to word frequency effects (Hulme et al., 1997), phonological similarity effects (Li, Schweickert, & Gandour 2000), irrelevant speech effects (Buchner & Erdfelder, 2005), and other factors affecting short-term memory such as serial position, word length, and lexicality (Schweickert, Chen, & Poirier, 1999).

Hindsight Bias

Memory for previous judgments is often biased by outcome knowledge attained later. This phenomenon has been called hindsight bias. MPT models of hindsight bias have typically been applied to the hindsight memory design: Participants first provide original numerical judgments on a set of items, for example, general knowledge questions. After receiving feedback on some of the items (experimental items) they are later asked to recall all their original judgments. Hindsight bias is observed when the recall for experimental items is biased toward the correct answer.

Basic Model

Erdfelder and Buchner (1998b) proposed an MPT model decomposing different processes leading to hindsight bias, namely reconstruction and recollection biases. Additionally, the model contains several parameters characterizing the distribution of responses as well as guessing parameters. To apply this model, continuous responses need to be categorized.

Extensions

Dehn and Erdfelder (1998) experimentally manipulated feedback answers and proposed a variant of the Erdfelder-Buchner model for analyzing their data.

Applications

Pohl and Gawlik (1995) suggested a slightly modified MPT model and applied it to both hindsight and misinformation effect experiments.

Using the Erdfelder and Buchner (1998b) model, Bayen, Erdfelder, Bearden, and Lozito (2006) studied the mechanisms underlying larger hindsight bias effects in older people. Finally, Erdfelder, Brandt, and Bröder (2007) further investigated how interference effects contribute to recollection bias and found evidence for both item-specific and item-unspecific retroactive inhibition in hindsight bias.

Other Memory Models

Space limitations preclude a discussion of additional MPT models that have been proposed for various learning and memory tasks as well as cognitive illusions. Numerous domains such as prospective memory (Smith & Bayen, 2004, 2005, 2006), knowledge assessment (Müller, 2004), neuropsychological assessment (Reiter, 2000), misinformation effects (Jacoby, Bishara, Hessels, & Toth, 2005), and the illusion of truth (Unkelbach & Stahl, 2009) are among the fields that have been addressed by MPT models.

Wason Selection Task

A classical reasoning task was developed by Wason (1966) and it has long been known as the Wason selection task (WST): Four cards are presented to participants, each bearing a letter on one side and a number on the other. The visible card sides read "A", "D", "3", and "7". Reasoners are asked to select those cards – and those cards only – which must be turned around to test the rule "if there is an A on one side of the card, then there is a 3 on the other side of the card".

Basic Model

A first stochastic model for the WST was presented by Evans (1977). He assumed that card selection is independent for each of the options and thus modeled the marginal probabilities through four parameters – each of which represented the probability of selecting a card. Additionally, Evans considered four different rules and proposed a combined MPT model with two parameters representing logical reasoning and two additional parameters capturing selections based on nonlogical evidence or response tendencies.

Extensions

Krauth (1982) suggested an extension of Evans' (1977) independence model to include three finite states, viz., verification, falsification, and matching tendencies. However, like Evans' original model only marginal probabilities of card selection were considered. By contrast, Klauer, Stahl, and Erdfelder (2007) recently proposed a hierarchy of models that refer to all $2^4 = 16$ possible card selection patterns in the WST. The inference-guessing model proved to be the most successful model of this family in a series of six validation experiments. It can be seen as a formalization of Evans and Over's (2004) dual-process theory of reasoning and comprises submodels representing logical reasoning (inference model) and independent card selections (implying heuristic processes or mere guessing). This model was recently extended by Stahl, Klauer, and Erdfelder (2008) to study the matching bias phenomenon.

Propositional Reasoning With Conditionals

A similar propositional reasoning task involves evaluations of four conditional inferences corresponding to (1) modus ponens, (2) acceptance of the consequent, (3) denial of the antecedent, and (4) modus tollens. Again, 2^4 response patterns are possible across these four tasks.

Basic Models

Building upon prior work by Klauer and Oberauer (1995), Oberauer (2006) formalized several different propositional reasoning theories – for example, the theory of mental models (Johnson-Laird & Byrne, 2002) and the dual-process (suppositional) theory (Evans & Over, 2004) – as MPT models. A modified mental-model theory and a dual-process model variant best accounted for his data.

Syllogistic Reasoning

Unlike propositional reasoning, syllogistic reasoning involves the use of quantifiers like “all”, “some”, and “none” in the reasoning process.

Basic Model

Klauer, Musch, and Naumer (1999) developed an MPT model to study belief bias in syllogistic reasoning. “Belief bias” refers to the phenomenon that – independent of logical correctness – plausible conclusions are more often accepted than implausible ones. Klauer et al. (1999) showed that this effect can be accommodated by a model including parameters that represent the probabilities of logical reasoning and plausibility assessments given failures of logical reasoning. Tests of syllogistic reasoning theories within this model framework showed that a modified mental-model theory

can best account for the results – a claim which has attracted substantial debate (Garnham & Oakhill, 2005; Klauer & Musch, 2005).

Perception Paradigms

MPT models have also been developed for object as well as letter recognition (Ashby, Prinzmetal, Ivry, & Maddox, 1996; Brown, 1998; Giesbrecht & Dixon, 1999; Hübner & Volberg, 2005; Prinzmetal, Ivry, Beck, & Shimizu, 2002), speech recognition (Batchelder & Crowther, 1997), word recognition (Maris, 2002; Ratcliff & McKoon, 2001), and negative priming experiments (Mayr & Buchner, 2006, in press; Mayr, Buchner, & Dentale, 2009; Mayr, Hauke, & Buchner, in press). These models proved to be useful in research on perception and selective attention.

MPT Models in Other Fields of Psychology

Applications of MPT models are not limited to cognitive psychology but have entered other branches of behavioral science as well. Most of these applications relate to problems of psychological assessment. We will briefly sketch the most influential of these developments.

Consensus Analysis

In many areas of psychology, assessment of individual differences in competency is pursued. Consensus analysis (Batchelder & Romney, 1986, 1988, 1989; Romney, Batchelder, & Weller, 1987; Romney, Weller, & Batchelder, 1986) aims at determining these differences in situations with unknown answer keys (i.e., unknown correct responses).

Basic Models

Romney et al. (1986) introduced the general condorcet model (GCM) as an extension of the 2HT model (Snodgrass & Corwin, 1988) outlined above. In its most general form, the GCM assumes heterogeneity of person abilities, response tendencies, and item difficulties for dichotomous responses. A respondent knows the correct response to an item based on his or her competency and the item's difficulty. If the correct response is unknown, the testee will guess. If homogeneous item difficulties are assumed, the GCM is identifiable and can be analyzed using the MPT framework.

Extensions

Klauer and Batchelder (1996, see also Klauer, 1996) used a model similar to the GCM to estimate inter-rater agreement

for categorization tasks (e.g., in a clinical setting, classifying patients as psychotic vs. neurotic). The competency parameter, however, reflects the amount of agreement between raters rather than individual competency. Consequently, this parameter can be interpreted as an inter-rater reliability measure.

Applications

Although the GCM found widespread use in anthropology, expertise research, and cross-cultural studies (cf. Romney & Batchelder, 1999) it has not yet been applied in cognitive psychology.

Multiple-Choice Tests

Multiple-choice tasks are frequently used across various social sciences and beyond. Interestingly, no sooner than the late 1980s, MPT models were introduced in this field.

Basic Models

Garcia-Perez (1990, 1993) and Garcia-Perez and Fray (1991) introduced a binary tree model which accounts not only for correct responses through conclusive knowledge or lucky guessing, but also for situations in which partial knowledge leads to a correct response (e.g., correctly ruling out response alternatives as lures).

Randomized Response Technique

It is well known that the validity of self-report data suffers from the tendency to bias responses in the direction of what is considered socially acceptable. The basic idea of the randomized response technique (RRT; Warner, 1965) is to add random noise to the responses such that the true status of an individual on the respective attribute is not identifiable on grounds of his or her response. Given that respondents are more likely to be honest when their true status cannot be inferred from their response, more honest responding is encouraged and has in fact been observed (for a meta-analysis, see Lensvelt-Mulders, Hox, van der Heijden, & Maas, 2005).

Basic Models

Several variants of the RRT have been proposed and successfully employed to obtain information about sensitive attributes (for a review, see Antonak & Livneh, 1995). In the forced-response variant of the RRT, a randomization device (e.g., the participant's month of birth) is used to determine whether participants are requested to respond honestly or to provide a prespecified response (e.g., "yes"). Because the distribution of the randomization device is known beforehand, the prevalence of the attribute under consider-

ation can be estimated from the observed proportion of *yes* responses. MPT models for the RRT require a parameter reflecting the prevalence of the sensitive attribute and a second parameter to absorb the proportion of prompted *yes* responses. Although ML estimates can be easily obtained as closed-form solutions, the MPT approach offers several benefits, including significance tests of constraints on the parameters, multiple group, and power analyses.

Extensions

An important extension of the forced-response variant is to assume that a certain proportion of respondents fail to comply with the randomized response instructions by denying to respond as requested by the outcome of the randomization device (Clark & Deshamais, 1998). The corresponding MPT model adds a parameter reflecting the proportion of noncompliant respondents and splits the sample into two parts with different probabilities of being asked to respond truthfully, leading to a joint MPT model with four parameters (Musch, Bröder, & Klauer, 2001; Ostapczuk, Moshagen, Zhao, & Musch, 2009).

Applications

The Clark and Deshamais (1998) model showed promise in various applications, for example, to investigate the education effect in stereotypes toward foreigners (Ostapczuk, Musch, & Moshagen, in press) or the prevalence of tax evasion (Musch et al., 2001).

Implicit Attitude Measurement

Recently, MPT models have also been introduced in the field of implicit attitude measurement.

Basic Model

According to the Quad model (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005) there are four basic processes underlying implicit attitude measurement tasks: (1) *association activation* represents the probability that a given stimulus triggers an automatic valence association; (2) *discriminability* corresponds to the probability that a given stimulus can be categorized correctly; (3) *overcoming bias* represents the conditional probability that an automatic association can be inhibited if it interferes with the discrimination process; and (4) *guessing*.

Extensions

Several studies have demonstrated an excellent Quad model fit for data of the implicit association test (IAT; Conrey et al., 2005; Greenwald, McGhee, & Schwartz, 1998; Sherman et al., 2008). However, the model is not suitable for all

implicit attitude paradigms (e.g., Stahl & Degner, 2007). Therefore, alternative MPT models have been proposed such as the ABC model for the extrinsic affective Simon task (De Houwer, 2003; Stahl & Degner, 2007; Stahl & Unkelbach, 2009), an extended Quad model variant for the Go/No-go association task (Gonsalkorale, von Hippel, Sherman, & Klauer, 2009; Nosek & Banaji, 2001), and different process dissociation models for the weapon identification task (Bishara & Payne, 2009; Klauer & Voss, 2008; Payne, 2001).

Applications

Several open questions in implicit attitude research have been answered by means of MPT models. For instance, it has been shown that age-related differences in IAT scores are rarely due to differences in automatic associations but rather to differences in the ability to inhibit such automatic associations (Gonsalkorale, Sherman, & Klauer, 2009; for a review of Quad model applications, see Sherman et al., 2008).

Developments in the Statistical Methodology of MPT Models

Three topics have dominated the methodological literature on MPT models in the past decade: (1) tests for small samples, (2) generalizations of the MPT model framework, and (3) heterogeneity between participants and/or items.

- (1) *Tests.* The statistical tests typically used to assess model fit or to compare parameters in MPT models are asymptotic tests that may produce misleading results in case of small samples. Several alternatives have been proposed for small samples, for example, exact tests (Garcia-Perez, 2000; Garcia-Perez & Nunez-Anton, 2004) and the parametric bootstrap (Efron & Tibshirani, 1997).
- (2) *Generalizations.* Hu (2001) proposed an extension of MPT models to (mean) response times. Moreover, nonstandard parameter constraints have been shown to be compatible with the MPT model structure, for example, order constraints (Baldi & Batchelder, 2003; Knapp & Batchelder, 2004) or log-linear main effects without interactions (Klauer et al., 1999). Recently, Schweickert and Chen (2008) showed that, given sufficiently large samples and experimental designs in which factors are known to affect model parameters selectively, the correct MPT model can in some cases be inferred from the empirical data structure.
- (3) *Heterogeneity.* Standard MPT analyses require homogeneous model parameters across participants and items.

This assumption is often implausible. Although standard MPT analyses of heterogeneous frequency data may sometimes be reasonable (Chechile, in press b), modeling parameter heterogeneity between participants and/or items was probably the hottest topic in recent methodological discussions. Several tests to assess parameter homogeneity versus heterogeneity have been proposed (Klauer, 2006; Smith & Batchelder, 2008). Also, different model frameworks have been suggested that can address parameter heterogeneity appropriately: Rasch model generalizations of MPT models (Batchelder, 1998), Bayesian MPT models based on different prior distributions of the person parameters (Klauer, in press; Smith & Batchelder, in press), and latent-class MPT models (Klauer, 2006). Rasch generalizations have been applied successfully to some simple MPT models (e.g., Maris, 1995) but have so far not been developed to handle any possible MPT model. By contrast, Klauer's (2006) hierarchical latent-class framework and Bayesian MPT models based on beta prior distributions (Smith & Batchelder, in press) or multivariate normal prior distributions (Klauer, in press) can be applied to any MPT model.

Computer Programs for MPT Models

Initiated by the pioneering work of Hu and Batchelder (1994), five major general-purpose MPT programs have been developed in the past 15 years: MBT (Hu, 1999), GPT (Hu & Phillips, 1999), AppleTree (Rothkegel, 1999), HMMTree (Stahl & Klauer, 2007), and multiTree (Moshagen, 2009).³

MBT, a DOS program that limits the size of a model to 60 branches, has largely been superseded by its successors both in terms of functionality and usability, and is therefore not considered in detail. GPT and HMMTree are Windows programs. AppleTree runs on MacOS, but requires the classic environment that has been discontinued with the advent of the new Mac OS X operation system "Leopard". multiTree is a java program running on Linux, MacOS, and Windows environments.

Each of these programs provides the basic statistical computations needed for MPT modeling, including parameter estimation and goodness-of-fit tests. GPT, AppleTree, and multiTree additionally allow users to change various aspects related to parameter estimation, such as the lambda value of the power divergence statistic, the stepsize parameter, and the criterion of convergence. All programs also provide some information on the variability of parameters and their variance-covariance matrix. multiTree can also be used to obtain bootstrapped standard errors and confidence intervals (Efron & Tibshirani, 1997).

The definition of MPT models proceeds by entering equations as a plain text file, with very similar formats for each of the programs. AppleTree, GPT, and multiTree

³ multiTree is available free of charge at <http://psycho3.uni-mannheim.de/multitree>.

additionally offer graphical input facilities to define and modify the trees of a model. Observed category frequencies are also provided in plain text files which are largely compatible across all programs. Except for HMMTree, each program can store and analyze multiple data files in a batch mode.

Apart from this basic functionality, the programs differ with respect to more advanced modeling features. HMMTree is the only program available that can be used to analyze latent-class hierarchical multinomial models, and only GPT can handle nonbinary multinomial models. Both AppleTree and GPT contain basic checks for identifiability, but multiTree additionally computes the rank of the Jacobian (Bamber & van Santen, 2000), and can determine simulated identifiability (Rouder & Batchelder, 1998). GPT and multiTree can also be used to conduct simulation studies, for example, to evaluate the robustness against violations of assumptions. Given the importance of statistical power, features related to power analyses are of great interest. Post hoc power analysis concerning global model misfit given the effect size metric w (Cohen, 1988) can be conducted with AppleTree and GPT. In addition, multiTree contains routines for performing a priori and post hoc power analyses for particular parameter restrictions, that is, effects defined by the difference between the parameters of a H_0 and a H_1 model (Erdfelder, Faul, & Buchner, 2005).

All of these general-purpose programs implement the EM algorithm tailored for MPT models, and thus cannot be used to analyze parametric models that do not belong to the family of MPT models. In some instances it may be helpful to rely on more general Newton-Raphson-type methods for parameter estimation. This can be achieved by using Microsoft Excel's SOLVER library with the macros provided by Dodson, Prinzmetal, and Shimamura (1998).

Acknowledgments

Manuscript preparation was supported by grants from the Deutsche Forschungsgemeinschaft (SFB 504, Project A12, and Mu 2674/1-1). The authors would like to thank Janina Hoffmann, Daniela Link, Kerstin Mertz, and Berenike Waubert de Puiseau for their help in preparing this review.

References

- Aleman, A., Böcker, K. B. E., Hijman, R., de Haan, E. H. F., & Kahn, R. S. (2003). Cognitive basis of hallucinations in schizophrenia: Role of top-down information processing. *Schizophrenia Research*, 64, 175–185.
- Antonak, R. F., & Livneh, H. (1995). Randomized response technique: A review and proposed extension to disability attitude research. *Genetic, Social, and General Psychology Monographs*, 121, 97–145.
- Ashby, F. G., Prinzmetal, W., Ivry, R., & Maddox, W. T. (1996). A formal theory of feature binding in object perception. *Psychological Review*, 103, 165–192.
- Baldi, P., & Batchelder, W. H. (2003). Bounds on variances of estimators for multinomial processing tree models. *Journal of Mathematical Psychology*, 47, 467–470.
- Bamber, D., & van Santen, J. P. H. (2000). How to assess a model's testability and identifiability. *Journal of Mathematical Psychology*, 44, 20–40.
- Batchelder, W. H. (1998). Multinomial processing tree models and psychological assessment. *Psychological Assessment*, 10, 331–344.
- Batchelder, W. H., Chosak-Reiter, J., Shankle, W. R., & Dick, M. B. (1997). A multinomial modeling analysis of memory deficits in Alzheimer's and vascular dementia. *Journal of Gerontology: Psychological Sciences and Social Sciences*, 52B, 206–215.
- Batchelder, W. H., & Crowther, C. S. (1997). Multinomial processing tree models of factorial categorization. *Journal of Mathematical Psychology*, 41, 45–55.
- Batchelder, W. H., & Riefer, D. M. (1980). Separation of storage and retrieval factors in free recall of clusterable pairs. *Psychological Review*, 87, 375–397.
- Batchelder, W. H., & Riefer, D. M. (1986). The statistical analysis of a model for storage and retrieval processes in human memory. *British Journal of Mathematical and Statistical Psychology*, 39, 129–149.
- 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.
- Batchelder, W. H., & Riefer, D. M. (2007). Using multinomial processing tree models to measure cognitive deficits in clinical populations. In R. W. J. Neufeld (Ed.), *Advances in clinical cognitive science: Formal modeling of processes and symptoms* (pp. 19–50). Washington, DC: APA Books.
- Batchelder, W. H., Riefer, D. M., & Hu, X. (1994). Measuring memory factors in source monitoring: Reply to Kinchla. *Psychological Review*, 101, 172–176.
- Batchelder, W. H., & Romney, A. K. (1986). The statistical analysis of a general condorcet model for dichotomous choice situations. In B. Grofman & G. Owen (Eds.), *Information pooling and group decision making* (pp. 103–112). Greenwich, CT: JAI Press.
- Batchelder, W. H., & Romney, A. K. (1988). Test theory without an answer key. *Psychometrika*, 53, 71–92.
- Batchelder, W. H., & Romney, A. K. (1989). New results in test theory without the answer key. In E. E. Roskam (Ed.), *Mathematical psychology in progress* (pp. 229–248). Berlin: Springer.
- Bayen, U. J., Erdfelder, E., Bearden, J. N., & Lozito, J. P. (2006). The interplay of memory and judgment processes in effects of aging on hindsight bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1003–1018.
- 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.
- Bell, R., & Buchner, A. (in press). Valence modulates source memory for faces. *Memory & Cognition*.
- Bell, R., Buchner, A., & Mund, I. (2008). Age-related differences in irrelevant-speech effects. *Psychology and Aging*, 23, 377–391.
- Bellezza, F. S. (2003). Evaluation of six multinomial models of conscious and unconscious processes with the recall-recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 779–796.

- Bender, R. H., Wallsten, T. S., & Ornstein, P. A. (1996). Age differences in encoding and retrieving details of a pediatric examination. *Psychonomic Bulletin & Review*, 3, 188–198.
- Bishara, A. J., & Payne, B. K. (2009). Multinomial process tree models of control and automaticity in weapon misidentification. *Journal of Experimental Social Psychology*, 45, 524–534.
- Brainerd, C. J., Reyna, V. F., & Mojardin, A. H. (1999). Conjoint recognition. *Psychological Review*, 106, 160–179.
- Bröder, A. (in press). Semantically clustered words are stored with integrated context. Validating a measurement model for source memory, storage, and retrieval in free recall. *Zeitschrift für Psychologie / Journal of Psychology*.
- Bröder, A., & Bredenkamp, J. (1996). SLIP-technique, process dissociation model, and multinomial modeling: New tools to experimentally prove Freudian slips? [Slip-Technik, Prozessdissoziationsmodell und multinomiale Modellierung: Neue Werkzeuge zum experimentellen Nachweis "Freud-scher Versprecher"?]. *Zeitschrift für Experimentelle Psychologie*, 43, 175–202.
- Bröder, A., Herwig, A., Teipel, S., & Fast, K. (2008). Different storage and retrieval deficits in normal aging and mild cognitive impairment: A multinomial modeling analysis. *Psychology and Aging*, 23, 353–365.
- Bröder, A., & Meiser, T. (2007). Measuring source memory. *Zeitschrift für Psychologie / Journal of Psychology*, 215, 52–60.
- Bröder, A., Noethen, D., Schütz, J., & Bay, P. (2007). Utilization of covariation knowledge in source monitoring: No evidence for implicit processes. *Psychological Research*, 71, 524–538.
- 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.
- Brown, V. (1998). Comparing parallel and sequential multinomial models of letter identification. In C. E. Dowling, F. S. Roberts, & P. Theuns (Eds.), *Recent progress in mathematical psychology* (pp. 253–284). Mahwah, NJ: Erlbaum.
- Buchner, A., Bell, R., Mehl, B., & Musch, J. (2009). No enhanced recognition memory, but better source memory for faces of cheaters. *Evolution and Human Behavior*, 30, 212–224.
- Buchner, A., & Erdfelder, E. (1996). On assumptions of, relations between, and evaluations of some process dissociation measurement models. *Consciousness and Cognition*, 5, 581–594.
- Buchner, A., & Erdfelder, E. (2005). Word frequency of irrelevant speech distractors affects serial recall. *Memory & Cognition*, 33, 86–97.
- Buchner, A., Erdfelder, E., Steffens, M., & Martensen, H. (1997). The nature of memory processes underlying recognition judgments in the process dissociation procedure. *Memory & Cognition*, 25, 508–517.
- Buchner, A., Erdfelder, E., & Vaterrodt-Plünnecke, B. (1995). Toward unbiased measurement of conscious and unconscious memory processes within the process dissociation framework. *Journal of Experimental Psychology: General*, 124, 137–160.
- Buchner, A., Steffens, M. C., Erdfelder, E., & Rothkegel, R. (1997). A multinomial model to assess fluency and recollection in a sequence learning task. *Quarterly Journal of Experimental Psychology*, 50A, 631–663.
- Buchner, A., Steffens, M. C., & Rothkegel, R. (1998). On the role of fragmentary knowledge in a sequence learning task. *Quarterly Journal of Experimental Psychology*, 51A, 251–281.
- Caldwell, J. I., & Masson, M. E. (2001). Conscious and unconscious influences of memory for object location. *Memory & Cognition*, 29, 285–295.
- Chechile, R. A. (1987). Trace susceptibility theory. *Journal of Experimental Psychology: General*, 116, 203–222.
- Chechile, R. A. (2004). New multinomial models for the Chechile-Meyer task: Comment. *Journal of Mathematical Psychology*, 48, 364–384.
- Chechile, R. A. (2007). A model-based storage-retrieval analysis of developmental dyslexia. In R. W. J. Neufeld (Ed.), *Advances in clinical cognitive sciences: Formal modeling of processes and symptoms* (pp. 51–79). Washington, DC: APA Books.
- Chechile, R. A. (in press a). Modeling storage and retrieval processes with clinical populations with applications examining alcohol-induced amnesia and Korsakoff amnesia. *Journal of Mathematical Psychology*.
- Chechile, R. A. (in press b). Pooling data versus averaging model fits for some prototypical multinomial processing tree models. *Journal of Mathematical Psychology*.
- Chechile, R. A., & Meyer, D. L. (1976). A Bayesian procedure for separately estimating storage and retrieval components of forgetting. *Journal of Mathematical Psychology*, 13, 269–295.
- Clark, S. J., & Desharnais, R. A. (1998). Honest answers to embarrassing questions: Detecting cheating in the randomized response model. *Psychological Methods*, 3, 160–168.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Mahwah: Erlbaum (Revised edition).
- Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The Quad model of implicit task performance. *Journal of Personality and Social Psychology*, 89, 469–487.
- Cüpper, L., & Erdfelder, E. (2004). Die Wortstammlänge beeinflusst kontrollierte, nicht aber automatische Gedächtnisprozesse im Wortstammerngänzungsparadigma [The length of word stems influences controlled but not automatic memory processes in the word stem completion paradigm]. *Zeitschrift für Psychologie / Journal of Psychology*, 212, 167–176.
- De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology*, 50, 77–85.
- Dehn, D., & Erdfelder, E. (1998). What kind of bias is hindsight bias? *Psychological Research*, 61, 135–146.
- Dodson, C. S. (2007). Retrieval-based illusory recollections: Why study-test contextual changes impair source memory. *Memory & Cognition*, 35, 1211–1221.
- Dodson, C. S., Darragh, J., & Williams, A. (2008). Stereotypes and retrieval-provoked illusory source recollections. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 460–477.
- Dodson, C. S., Holland, P. W., & Shimamura, A. P. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1121–1136.
- Dodson, C. S., Prinzmetal, W., & Shimamura, A. P. (1998). Using Excel to estimate parameters from observed data: An example from source memory data. *Behavior Research Methods, Instruments, and Computers*, 30, 517–526.
- Efron, B., & Tibshirani, R. J. (1997). *An introduction to the bootstrap*. New York: Chapman & Hall.
- Ehrenberg, K., & Klauer, K. C. (2005). Flexible use of source information: Processing components of the inconsistency effect in person memory. *Journal of Experimental Social Psychology*, 41, 369–387.
- Elandt-Johnson, R. C. (1971). *Probability models and statistical methods in genetics*. New York: Wiley.
- Erdfelder, E., & Bayen, U. J. (1991). Episodisches Gedächtnis im Alter: Methodologische und empirische Argumente für einen Zugang über mathematische Modelle [Episodic memory and

- aging: Methodological and empirical arguments for an approach based on mathematical models]. In D. Frey (Ed.), *Bericht über den 37. Kongress der Deutschen Gesellschaft für Psychologie in Kiel 1990* (Vol. 2, pp. 172–180). Göttingen, Germany: Hogrefe.
- Erdfelder, E., Brandt, M., & Bröder, A. (2007). Recollection biases in hindsight judgments. *Social Cognition*, 25, 114–131.
- Erdfelder, E., & Bredenkamp, J. (1998). Recognition of script-typical and script-atypical information: Effects of cognitive elaboration. *Memory & Cognition*, 26, 922–938.
- Erdfelder, E., & Buchner, A. (1998a). Process-dissociation measurement models: Threshold theory or detection theory? *Journal of Experimental Psychology: General*, 127, 83–97.
- Erdfelder, E., & Buchner, A. (1998b). Decomposing the hindsight bias: A multinomial processing tree model for separating recollection and reconstruction in hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 387–414.
- Erdfelder, E., & Buchner, A. (2003). Prozessdissoziationsprozedur: Quo vadis? [Process dissociation procedure: Quo vadis?]. *Zeitschrift für Psychologie / Journal of Psychology*, 211, 17–25.
- 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.
- Erdfelder, E., Faul, F., & Buchner, A. (2005). Power analysis for categorical methods. In B. Everitt & D. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (pp. 1565–1570). London: Wiley.
- Evans, J. S. B. T. (1977). Toward a statistical theory of reasoning. *Quarterly Journal of Experimental Psychology*, 29, 621–635.
- Evans, J. S. B. T., & Over, D. E. (2004). *If*. Oxford, UK: Oxford University Press.
- Garcia-Perez, M. A. (1990). A comparison of two models of performance in objective tests: Finite states versus continuous distributions. *British Journal of Mathematical and Statistical Psychology*, 43, 73–91.
- Garcia-Perez, M. A. (1993). In defence of “none of the above”. *British Journal of Mathematical and Statistical Psychology*, 46, 213–229.
- Garcia-Perez, M. A. (2000). Exact finite-sample significance and confidence regions for goodness-of-fit statistics in one-way multinomials. *British Journal of Mathematical and Statistical Psychology*, 53, 193–207.
- Garcia-Perez, M. A., & Frary, R. B. (1991). Finite state polynomial item characteristic curves. *British Journal of Mathematical and Statistical Psychology*, 44, 45–73.
- Garcia-Perez, M. A., & Nunez-Anton, V. (2004). On the chi-square approximation to the exact distribution of goodness-of-fit statistics in multinomial models with composite hypotheses. *British Journal of Mathematical and Statistical Psychology*, 57, 73–96.
- Garnham, A., & Oakhill, J. V. (2005). Accounting for belief bias in a mental model framework: Comment on Klauer, Musch, and Naumer (2000). *Psychological Review*, 112, 509–518.
- Gawronski, B., Ehrenberg, K., Banse, R., Zukova, J., & Klauer, K. C. (2003). It's in the mind of the beholder: The impact of stereotypic associations on category-based and individuating impression formation. *Journal of Experimental Social Psychology*, 39, 16–30.
- Giesbrecht, B., & Dixon, P. (1999). Isolating the interference caused by cue duration in partial report: A quantitative approach. *Memory & Cognition*, 27, 220–233.
- Golz, D., & Erdfelder, E. (2004). Effekte von L-Dopa auf die Speicherung und den Abruf verbaler Informationen bei Schlaganfallpatienten [Effects of L-Dopa on storage and retrieval of verbal information in stroke patients]. *Zeitschrift für Neuropsychologie*, 15, 275–286.
- Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology*, 45, 410–414.
- Gonsalkorale, K., von Hippel, W., Sherman, J. W., & Klauer, K. C. (2009). Bias and regulation of bias in intergroup interactions: Implicit attitudes toward muslims and interaction quality. *Journal of Experimental Social Psychology*, 45, 161–166.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Hillstrom, A. P., & Logan, G. D. (1997). Process dissociation, cognitive architecture, and response time: Comments of Lindsay and Jacoby (1994). *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1561–1578.
- Hu, X. (1999). Multinomial processing tree models: An implementation. *Behavior Research Methods, Instruments, & Computers*, 31, 689–695.
- Hu, X. (2001). Extending general processing tree models to analyze reaction time experiments. *Journal of Mathematical Psychology*, 45, 603–634.
- Hu, X., & Batchelder, W. H. (1994). The statistical analysis of general processing tree models with the EM algorithm. *Psychometrika*, 59, 21–47.
- Hu, X., & Phillips, G. A. (1999). GPT.EXE: A powerful tool for the visualization and analysis of general processing tree models. *Behavior Research Methods, Instruments, & Computers*, 31, 220–234.
- Hübner, R., & Volberg, G. (2005). The integration of object levels and their content: A theory of global/local processing and related hemisphere differences. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 520–541.
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a reintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1217–1232.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30, 513–541.
- Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1–36.
- Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General*, 134, 131–148.
- Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, 109, 646–678.
- Keefe, R. S. E., Arnold, M. C., Bayen, U. J., & Harvey, P. D. (1999). Source monitoring deficits in patients with schizophrenia: A multinomial modeling analysis. *Psychological Medicine*, 29, 903–914.
- Keefe, R. S. E., Arnold, M. C., Bayen, U. J., McEvoy, J. P., & Wilson, W. H. (2002). Source monitoring deficits for self-generated stimuli in schizophrenia: Multinomial modeling of data from three sources. *Schizophrenia Research*, 57, 51–67.

- Kinchla, R. A. (1994). Comments on Batchelder and Riefer's multinomial model for source monitoring. *Psychological Review*, 101, 166–171.
- Klauer, K. C. (1996). Urteilerübereinstimmung für dichotome Kategoriensysteme [Interrater agreement for dichotomous categories]. *Diagnostica*, 42, 101–118.
- Klauer, K. C. (2006). Hierarchical multinomial processing tree models: A latent-class approach. *Psychometrika*, 71, 7–31.
- Klauer, K. C. (in press). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika*.
- Klauer, K. C., & Batchelder, W. H. (1996). Structural analysis of subjective categorical data. *Psychometrika*, 61, 199–239.
- Klauer, K. C., & Ehrenberg, K. (2005). Social categorization and fit detection under cognitive load: Efficient or effortful? *European Journal of Social Psychology*, 35, 493–516.
- Klauer, K. C., Ehrenberg, K., & Wegener, I. (2003). Crossed categorization and stereotyping: Structural analyses, effect patterns, and dissociative effects of context relevance. *Journal of Experimental Social Psychology*, 39, 332–354.
- Klauer, K. C., & Meiser, T. (2000). A source-monitoring analysis of illusory correlations. *Personality and Social Psychology Bulletin*, 26, 1074–1093.
- Klauer, K. C., & Musch, J. (2005). Accounting for belief bias in a mental model framework? No problem! Reply to Garnham and Oakhill (2005). *Psychological Review*, 112, 519–520.
- Klauer, K. C., Musch, J., & Naumer, B. (1999). On belief bias in syllogistic reasoning. *Psychological Review*, 107, 852–884.
- Klauer, K. C., & Oberauer, K. (1995). Testing the mental model theory of propositional reasoning. *Quarterly Journal of Experimental Psychology*, 48A, 671–687.
- Klauer, K. C., Stahl, C., & Erdfelder, E. (2007). The abstract selection task: New data and an almost comprehensive model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 680–703.
- Klauer, K. C., & Voss, A. (2008). Effects of race on responses and response latencies in the weapon identification task: A test of six models. *Personality and Social Psychology Bulletin*, 34, 1124–1140.
- Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the “Who Said What?” paradigm. *Journal of Personality and Social Psychology*, 75, 1155–1178.
- Klauer, K. C., Wegener, I., & Ehrenberg, K. (2002). Perceiving minority members as individuals: The effects of relative group size in social categorization. *European Journal of Social Psychology*, 32, 223–245.
- Knapp, B. R., & Batchelder, W. H. (2004). Representing parametric order constraints in multi-trial applications of multinomial processing tree models. *Journal of Mathematical Psychology*, 48, 215–229.
- Krantz, D. H. (1969). Threshold theories of signal detection. *Psychological Review*, 76, 308–324.
- Krauth, J. (1982). Formulation and experimental verification of models in propositional reasoning. *Quarterly Journal of Experimental Psychology*, 34A, 285–298.
- Lensvelt-Mulders, G. J. L. M., Hox, J. J., van der Heijden, P. G. M., & Maas, C. J. M. (2005). Meta-analysis of randomized response research: Thirty-five years of validation. *Sociological Methods & Research*, 33, 319–348.
- Li, X., Schweickert, R., & Gandour, J. (2000). The phonological similarity effect in immediate recall: Positions of shared phonemes. *Memory & Cognition*, 28, 1116–1125.
- Luce, R. D. (1963). A threshold theory for simple detection experiments. *Psychological Review*, 70, 61–79.
- Malmberg, K. J. (2002). On the form of ROCs constructed from confidence ratings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 380–387.
- Maris, E. (1995). Psychometric latent response models. *Psychometrika*, 60, 523–547.
- Maris, E. (2002). The role of orthographic and phonological codes in the word and the pseudoword superiority effect: An analysis by means of multinomial processing tree models. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1409–1431.
- Mayr, S., & Buchner, A. (2006). Evidence for episodic retrieval of inadequate prime responses in auditory negative priming. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 932–943.
- Mayr, S., & Buchner, A. (in press). Auditory negative priming endures response modality change, prime-response retrieval does not. *Quarterly Journal of Experimental Psychology*.
- Mayr, S., Buchner, A., & Dentale, S. (2009). Prime retrieval of motor responses in negative priming. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 408–423.
- Mayr, S., Hauke, R., & Buchner, A. (in press). Auditory location negative priming: A case of feature mismatch. *Psychonomic Bulletin & Review*.
- McBride, D. M., & Doshier, B. A. (1999). Forgetting rates are comparable in conscious and automatic memory: A process-dissociation study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 583–607.
- McBride, D. M., Doshier, B. A., & Gage, N. M. (2001). A comparison of forgetting for conscious and automatic memory processes in word fragment completion tasks. *Journal of Memory and Language*, 45, 585–615.
- McBride, D. M., & Shouder, H. (2003). Conceptual processing effects on automatic memory. *Memory & Cognition*, 31, 393–400.
- Meiser, T. (2000). Gedächtnisvorteil ausgeführter Handlungen: Anmerkungen zur multinomialen Gedächtnistheorie [The advantage of performed acts: Notes on multinomial memory theory]. *Zeitschrift für Experimentelle Psychologie*, 47, 1–12.
- Meiser, T. (2005). A hierarchy of multinomial models for multidimensional source monitoring. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 1, 2–17.
- Meiser, T., & Bröder, A. (2002). Memory for multidimensional source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 116–137.
- Meiser, T., & Hewstone, M. (2004). Cognitive processes in stereotype formation: The role of correct contingency learning for biased group judgments. *Journal of Personality and Social Psychology*, 87, 599–614.
- Meiser, T., & Hewstone, M. (2006). Illusory and spurious correlations: Distinct phenomena or joint outcomes of exemplar-based category learning? *European Journal of Social Psychology*, 36, 315–336.
- Meiser, T., & Sattler, C. (2007). Boundaries of the relation between conscious recollection and source memory for perceptual details. *Consciousness and Cognition*, 16, 189–210.
- Meiser, T., Sattler, C., & von Hecker, U. (2007). Metacognitive inferences in source memory judgments: 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.
- Menon, M., & Woodward, T. S. (2007). A comparison of one-high-threshold and two-high-threshold multinomial models of source monitoring. *Journal of Modern Applied Statistical Methods*, 6, 279–290.

- Miller, E., & Lewis, P. (1977). Recognition memory in elderly patients with depression and dementia: A signal detection analysis. *Journal of Abnormal Psychology*, 86, 84–86.
- Moshagen, M. (2009). multiTree: A computer program for the analysis of multinomial processing tree models. Manuscript submitted for publication.
- Müller, B. (2004). A multinomial model to assess central characteristics of mental operators. *Experimental Psychology*, 51, 201–213.
- Musch, J., Bröder, A., & Klauer, K. C. (2001). Improving survey research on the world-wide web using the randomized response technique. In U. Reips & M. Bosnjak (Eds.), *Dimensions of Internet science* (pp. 172–192). Lengerich, Germany: Pabst.
- Nosek, B. A., & Banaji, M. R. (2001). The Go/No-go association task. *Social Cognition*, 19, 625–666.
- Oberauer, K. (2006). Reasoning with conditionals: A test of formal models of four theories. *Cognitive Psychology*, 53, 238–283.
- Ostapczuk, M., Moshagen, M., Zhao, Z., & Musch, J. (2009). Assessing sensitive attributes using the randomized-response-technique: Evidence for the importance of response symmetry. *Journal of Educational and Behavioral Statistics*, 34, 267–287.
- Ostapczuk, M., Musch, J., & Moshagen, M. (in press). A randomized-response investigation of the education effect in attitudes towards foreigners. *European Journal of Social Psychology*.
- Ott, R., Curio, I., & Scholz, O. B. (2000). Implicit memory to auditorily presented threatening stimuli – A process-dissociation approach. *Perceptual and Motor Skills*, 90, 131–146.
- Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, 81, 181–192.
- Pohl, R. F., & Gawlik, B. (1995). Hindsight bias and the misinformation effect: Separating blended recollections from other recollection types. *Memory*, 3, 21–55.
- Prinzmetal, W., Ivry, R. B., Beck, D., & Shimizu, N. (2002). A measurement theory of illusory conjunctions. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 251–269.
- Ratcliff, R., & McKoon, G. (2001). A multinomial model for short-term priming in word identification. *Psychological Review*, 108, 835–846.
- Read, T. R. C., & Cressie, N. A. C. (1988). *Goodness-of-fit statistics for discrete multivariate data*. New York: Springer.
- Reiter, J. C. (2000). Measuring cognitive processes underlying picture naming in Alzheimer's and cerebrovascular dementia: A general processing tree approach. *Journal of Clinical and Experimental Neuropsychology*, 22, 351–369.
- Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, 95, 318–339.
- Riefer, D. M., & Batchelder, W. H. (1991). Age differences in storage and retrieval: A multinomial modeling analysis. *Bulletin of the Psychonomic Society*, 29, 415–418.
- Riefer, D. M., & Batchelder, W. H. (1995). A multinomial modeling analysis of the recognition-failure paradigm. *Memory & Cognition*, 23, 611–630.
- Riefer, D. M., Chien, Y., & Reimer, J. F. (2007). Positive and negative generation effects in source monitoring. *Quarterly Journal of Experimental Psychology*, 60, 1389–1405.
- Riefer, D. M., Hu, X., & Batchelder, W. H. (1994). Response strategies in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 680–693.
- Riefer, D. M., Knapp, B. R., Batchelder, W. H., Bamber, D., & Manifold, V. (2002). Cognitive psychometrics: Assessing storage and retrieval deficits in special populations with multinomial processing tree models. *Psychological Assessment*, 14, 184–201.
- Riefer, D. M., & LaMay, M. L. (1998). Memory for common and bizarre stimuli: A storage-retrieval analysis. *Psychonomic Bulletin & Review*, 5, 312–317.
- Riefer, D. M., & Rouder, J. N. (1992). A multinomial modeling analysis of the mnemonic benefits of bizarre imagery. *Memory & Cognition*, 20, 601–611.
- Romney, A. K., & Batchelder, W. H. (1999). Cultural consensus theory. In R. A. Wilson & F. C. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 208–209). Cambridge, MA: The MIT press.
- Romney, A. K., Batchelder, W. H., & Weller, S. C. (1987). Recent applications of cultural consensus theory. *American Behavioral Scientist*, 31, 163–177.
- Romney, A. K., Weller, S. C., & Batchelder, W. H. (1986). Culture as consensus – a theory of culture and informant accuracy. *American Anthropologist*, 88, 313–338.
- Rothkegel, R. (1999). AppleTree: A multinomial processing tree modeling program for Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, 31, 696–700.
- Rouder, J. N., & Batchelder, W. H. (1998). Multinomial models for measuring storage and retrieval processes in paired associate learning. In C. E. Dowling, F. S. Roberts, & P. Theuns (Eds.), *Recent progress in mathematical psychology* (pp. 195–225). Mahwah, NJ: Erlbaum.
- Rouder, J. N., Lu, J., Morey, R. D., Sun, D., & Speckman, D. L. (2008). A hierarchical process-dissociation model. *Journal of Experimental Psychology: General*, 137, 370–389.
- Schmitter-Edgecombe, M., Marks, W., Wright, M. J., & Ventura, M. (2004). Retrieval inhibition in directed forgetting following severe close-head injury. *Neuropsychology*, 18, 104–114.
- Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition*, 21, 168–175.
- Schweickert, R., & Chen, S. (2008). Tree inference with factors selectively influencing processes in a processing tree. *Journal of Mathematical Psychology*, 52, 158–183.
- Schweickert, R., Chen, S., & Poirier, M. (1999). Redintegration and the useful lifetime of the verbal memory representation. *International Journal of Psychology*, 34, 447–453.
- Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review*, 115, 314–335.
- Sherman, J. W., Groom, C. J., Ehrenberg, K., & Klauer, K. C. (2003). Bearing false witness under pressure: Implicit and explicit components of stereotype-driven memory distortions. *Social Cognition*, 21, 213–246.
- Simons, J. S., Dodson, C. S., Bell, D., & Schacter, D. L. (2004). Specific- and partial-source memory: Effects of aging. *Psychology and Aging*, 19, 689–694.
- Simons, J. S., Verfaellie, M., Galton, C. J., Miller, B. L., Hodges, J. R., & Graham, K. S. (2002). Recollection-based memory in frontotemporal dementia: Implications for theories of long-term memory. *Brain*, 125, 2523–2536.
- Smith, J. B., & Batchelder, W. H. (2008). Assessing individual differences in categorical data. *Psychonomic Bulletin & Review*, 15, 713–731.
- Smith, J. B., & Batchelder, W. H. (in press). Beta-MPT: Multinomial processing tree models for addressing individual differences. *Journal of Mathematical Psychology*.
- Smith, R. E., & Bayen, U. J. (2004). A multinomial model of event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 756–777.
- Smith, R. E., & Bayen, U. J. (2005). The effects of working memory resource availability on prospective memory: A

- formal modeling approach. *Experimental Psychology*, 52, 243–256.
- Smith, R. E., & Bayen, U. J. (2006). The source of adult age differences in event-based prospective memory: A multinomial modeling approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 623–635.
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117, 34–50.
- Spaniol, J., & Bayen, U. J. (2002). When is schematic knowledge used in source monitoring? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 631–651.
- Stahl, C. (2006). Multinomiale Verarbeitungsbaummodelle in der Sozialpsychologie. *Zeitschrift für Sozialpsychologie*, 37, 161–171.
- Stahl, C., & Degner, J. (2007). Assessing automatic activation of valence. A multinomial model of EAST performance. *Experimental Psychology*, 54, 99–112.
- Stahl, C., & Klauer, K. C. (2007). HMMTree: A computer program for latent-class hierarchical multinomial processing tree models. *Behavior Research Methods*, 39, 267–273.
- 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.
- Stahl, C., & Klauer, K. C. (2009). Measuring phantom recollection in the simplified conjoint recognition paradigm. *Journal of Memory and Language*, 60, 180–193.
- Stahl, C., Klauer, K. C., & Erdfelder, E. (2008). Matching bias in the selection task is not eliminated by explicit negations. *Thinking & Reasoning*, 14, 281–303.
- Stahl, C., & Unkelbach, C. (2009). Evaluative learning with single versus multiple unconditioned stimuli: The role of contingency awareness. *Journal of Experimental Psychology: Animal Behavior Processes*, 35, 286–291.
- Steffens, M. C., Buchner, A., Martensen, H., & Erdfelder, E. (2000). Further evidence on the similarity of memory processes in the process dissociation procedure and in source monitoring. *Memory & Cognition*, 28, 1152–1164.
- Steffens, M. C., Jelenec, P., & Mecklenbräuker, S. (2009). Decomposing the memory processes contributing to enactment effects by multinomial modeling. *European Journal of Cognitive Psychology*, 21, 61–83.
- Steffens, M. C., Jelenec, P., Mecklenbräuker, S., & Thompson, E. M. (2006). Decomposing retrieval and integration in memory for actions. A multinomial modeling approach. *Quarterly Journal of Experimental Psychology*, 59, 557–576.
- Stenberg, G. (2006). Conceptual and perceptual factors in the picture superiority effect. *European Journal of Cognitive Psychology*, 18, 813–847.
- Stolz, J. A., & Merikle, P. A. (2000). Conscious and unconscious influences of memory: Temporal dynamics. *Memory*, 8, 333–343.
- Strack, F., & Bless, H. (1994). Memory for non-occurrences: Metacognitive and presuppositional strategies. *Journal of Memory and Language*, 33, 203–217.
- Swets, J. A. (1961). Is there a sensory threshold? *Science*, 134, 168–177.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352–373.
- Tulving, E., & Wiseman, S. (1975). Relation between recognition and recognition failure of recallable words. *Bulletin of the Psychonomic Society*, 6, 79–82.
- Ulrich, R. (2009). Uncovering unobservable cognitive mechanisms: The contribution of mathematical models. In F. Rösler, C. Ranganath, B. Röder, & R. H. Kluwe (Eds.), *Neuroimaging of human memory: Linking cognitive processes to neural systems* (pp. 25–41). New York: Oxford University Press.
- Unkelbach, C., & Stahl, C. (2009). A multinomial modeling approach to dissociate different components of the truth effect. *Consciousness and Cognition*, 18, 22–38.
- Vaterrodt-Plünnecke, B., Krüger, T., & Bredenkamp, J. (2002). Process-dissociation procedure: A testable model for considering assumptions about the stochastic relation between consciously controlled and automatic processes. *Experimental Psychology*, 49, 3–26.
- Vaterrodt-Plünnecke, B., Krüger, T., Gerdes, H., & Bredenkamp, J. (1996). Prozess-Dissoziations-Prozedur: Prüfbare Messmodelle zur Erfassung von kontrollierten, automatischen und Antworttendenz-Prozessen [Process-dissociation procedure: Testable models to measure controlled, automatic and bias processes]. *Zeitschrift für Experimentelle Psychologie*, 43, 483–519.
- Vogt, V., & Brüder, A. (2007). Independent retrieval of source dimensions: An extension of results by Starns and Hicks (2005) and a comment on the ACSIM measure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 443–450.
- von Hecker, U., & Meiser, T. (2005). Defocused attention in depressed mood: Evidence from source monitoring. *Emotion*, 5, 456–463.
- Vorberg, D., & Schmidt, R. (1975). A testable property of some discrete-state models for recognition memory. *Psychological Review*, 82, 316–324.
- Warner, S. (1965). Randomized-response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60, 63–69.
- Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology I*. Harmondsworth, UK: Penguin.
- Wegener, I., De Beer, K., Schilling, G., Conrad, R., Imbierowicz, K., Geiser, F., et al. (2008). Patients with obesity show reduced memory for others' body shape. *Appetite*, 50, 359–366.
- Wegener, I., & Klauer, K. C. (2004). Inter-category versus intra-category fit: When social categories match social context. *European Journal of Social Psychology*, 34, 567–593.
- Wegener, I., & Klauer, K. C. (2005). Social categorization without fit: The effect of gender categories on the encoding of senseless syllables. *Zeitschrift für Sozialpsychologie*, 36, 91–101.
- Woodward, T. S., Menon, M., Hu, X., & Keefe, R. S. E. (2006). Optimization of a multinomial model for investigating hallucinations and delusions with source monitoring. *Schizophrenia Research*, 85, 106–112.
- Xu, M., & Bellezza, F. (2001). A comparison of the multitemory and detection theories of know and remember recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1197–1210.
- Yonelinas, A. P., & Jacoby, L. L. (1996). Response bias and the process dissociation procedure. *Journal of Experimental Psychology: General*, 125, 422–434.

Dr. Edgar Erdfelder

Lehrstuhl für Psychologie III
Universität Mannheim
Schloss Ehrenhof Ost 255
68131 Mannheim
Germany
Tel. +49 621/181-2146
Fax +49 621/181-3997
E-mail erdfelder@psychologie.uni-mannheim.de