Journal of Mathematical Psychology 73 (2016) 1–11

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Thou shalt identify! The identifiability of two high-threshold models in confidence-rating recognition (and super-recognition) paradigms

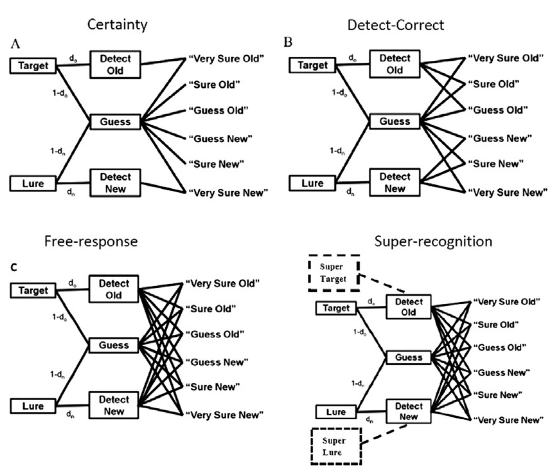
Rani Moran

School of Psychological Sciences, Tel Aviv University, Ramat Aviv, POB 39040, Tel Aviv, 69978, Israel

HIGHLIGHTS

- I examined the identifiability of 2HTM in confidence-rating recognition paradigms.
- I prove that the certainty and the detect–correct variants are identifiable.
- I prove that the free response mapping (FRM) variant is non- identifiable.
- Super-recognition, a novel paradigm with super-strong probes, identifies FRM.
- Solutions for the non-identifiability problem are discussed.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 21 November 2015
Received in revised form
23 January 2016

Keywords:
Identifiability proofs
Recognition memory
Model identifiability
Super-recognition
High thresholds
Computational models
Model comparison

ABSTRACT

Model identifiability i.e., the possibility to determine the 'true' parameters of a model in a unique manner based on an empirical dataset, is a vital property of any scientific model. Indeed, the absence of this property undermines the model's use qua measurement and inference tool. Here, I examined the identifiability of different variants of the successful two-high threshold model (2HTM) in the confidence-rating recognition paradigm. Traditionally, 2HTM has adopted the certainty assumption according to which, detected test-probes receive only the highest correct confidence rating. Modern variants, however, allow responses for detected items to distribute across all correct confidence levels (Detect–Correct; DC) or even across erroneous confidence levels (Free Response Mapping, FRM). Here, I present identifiability proofs for the certainty variant and, when there are multiple target conditions, for DC. Additionally, I present non-identifiability proofs for DC when there is a single target condition and for FRM. One important advantage of identifiability proofs over fitting-based methods for testing identifiability is that they highlight the mathematical principles that are instrumental for a deeper understanding of models. Based on these principles, I present a novel extended *super-confidence* paradigm, which includes super-strong targets and lures and that identifies FRM. I illustrate the perils of non-identifiability by reanalyzing a recent model comparison study of Chen et al. (2015). Finally, solutions for non-identifiability problems are discussed. While the current study focuses on identifiability in the context of recognition-memory, it should serve as a universal reminder for the importance of identifying models in any domain of research.

E-mail address: rani.moran@gmail.com.

http://dx.doi.org/10.1016/j.jmp.2016.03.002
0022-2496/

1. Introduction

Model *Identifiability* pertains to the theoretical possibility to determine (i.e., 'identify') in a unique manner the 'true' hidden parameters of a model, based on a dataset. This property is vital for any model—a perquisite for using it as an efficient research tool. For example, identifiability is a necessary condition for precise model-based statistical inferences. The other side of the coin is that non-identifiability entails several hazardous consequences such as, undermining a model's use qua measurement and inference tool and subjecting the model to inflated model-comparison penalties. Thus, non-identifiable models may underlie non-valid conclusions, leading research astray. Accordingly, the issue of model-identifiability has attracted substantial interest across diverse scientific domains such as statistics (e.g., Casella & Berger, 2002; Lehmann & Casella, 1998), economics (e.g., Fisher, 1966; Reiersøl, 1950) and psychology (e.g. Brainerd, Howe, & Desrochers, 1982; Brainerd, Howe, & Kingma, 1982). In the current paper, I examined the identifiability properties of the influential Two High Threshold Model (henceforth, 2HTM; Bayen, Murnane, & Erdfelder, 1996; Bröder & Meiser, 2007; Bröder & Schütz, 2009; Erdfelder & Buchner, 1998; Meiser, 2005; Meiser & Bröder, 2002; Province & Rouder, 2012; Snodgrass & Corwin, 1988) in the popular confidence-rating recognition paradigm.

Recognition memory is our ability to discriminate between 'old' objects that have been encountered in a given context and 'new' objects that have not. In laboratory settings, participants first study a list of items and are later tested on a different list, which consists of both studied (target) and unstudied (lure) items. Participants decide whether each test probe is 'old' or 'new'. Often, as in the paradigm focal to the current study, the old–new judgment is graded along a confidence scale, ranging from high confidence in 'old' to high confidence in 'new'.

According to 2HTM, one of the most successful recognition models, the output of the memory system is discrete (Pazzaglia, Dube, & Rotello, 2013). Specifically, targets and lures, respectively, can be either detected as old (with probability d_o) or as new (with probability d_n) or not-detected (see Fig. 1). Notably, both target and lure-detection are *high threshold* processes, meaning that lures are never detected as old (i.e., never pass the detect-old threshold) and targets are never detected as new. In case of non-detection the participant enters a state of uncertainty, i.e., a 'guessing' state.

Considering first the guessing state, it does not discriminate between targets and lures. Consequently, the probability of selecting a particular (confidence) response from this state is independent of the type of the test probe, target or lure—an instance of the *conditional independence* assumption (Province & Rouder, 2012). Additionally, guessed items can be rated with any of the old or new confidence judgments, with probabilities that are given by free, response-mapping, parameters.

Unlike the guessing state, the various 2HTM variants make different assumptions with respect to the confidence mappings from detect states. Traditionally, 2HTM has adopted the *certainty* assumption, according to which detect states always produce a correct response at the highest confidence level (Fig. 1(A)). This variant predicts linear ROC functions (Egan, 1958; Swets, 1986; Swets, Tanner, & Birdsall, 1964, Ch. 1). More recently, however, the certainty assumption has been critiqued as untenable and it has been argued instead (see Fig. 1(B)) that in the detect-states participants may distribute their responses across high and low confidence levels (e.g., Broadbent, 1966; Erdfelder & Buchner, 1998; Malmberg, 2002; Province & Rouder, 2012; Swagman, Province, & Rouder, 2015). Under this relaxed assumption, 2HTM can produce curved ROC functions (see Malmberg, 2002 for a detailed explanation)—meeting the challenge of accounting for typical empirical confidence-rating ROCs (e.g., Wixted, 2007). Notably, the model displayed in Fig. 1(B) (henceforth, DC) adheres to the *Detect–Correct* assumption according to which, detected items always yield a correct response: detected targets and lures are always assigned to one of the old and new confidence ratings, respectively.

Recently, Chen, Starns, and Rotello (2015) considered a more flexible '*Full Response Mapping*' 2HTM variant (henceforth FRM; Fig. 1(C)), which has also been presented elsewhere (e.g., Province & Rouder, 2012, Fig. 1(D); Swagman et al., 2015, Fig. 1(D)). In FRM, the detect–correct assumption is forsaken; any confidence rating can be given from the detect states, including erroneous ratings (e.g., new ratings for detected targets). Presumably, such erroneous ratings reflect response-errors and/or demand characteristics e.g., the common instructions to use the entire confidence scale. In sum, before FRM was introduced, 2HTM models were not allowed to make erroneous responses for detected items. Such responses were possible only in *low threshold* models (e.g., Krantz, 1969; Luce, 1963). For example, in a low threshold model a lure may be mistakenly detected as a target and consequently, be given an erroneous old response.

Here, I examined the identifiability of the various 2HTM variants. I begin with a definition of identifiability and I highlight the perils imposed by its absence. I prove that: (1) 2HTM with the certainty assumption is identifiable, (2) DC is identifiable if there are multiple target-strength conditions but non-identifiable if there is a single target condition, and (3) FRM is not identifiable. Such proofs comprise a valuable method for establishing model-identifiability or its absence, as they bring to the foreground mathematical principles that can aid researchers in developing a deeper and broader understanding for the structure and operation of models. Indeed, a close examination of the non-identifiability proof for FRM, revealed *why* this model is non-identifiable. This understanding inspired a novel extended paradigm, super-recognition, which renders the FRM identifiable. Next, revisiting a model comparison study of Chen et al. (2015), I illustrate how non-identifiable models can distort scientific conclusions. This investigation highlights the need to identify models and/or to devise appropriate solutions for the problem of non-identifiability. Such solutions are discussed in the final part of the paper.

2. Model identifiability

2.1. Establishing model-identifiability: theory and practice

A model is *globally identifiable*[1] if the mapping between model-parameters and model-predictions (also termed model-*outcomes*) is one to one over the entire parameter space. In other words, different combinations of parameters do not yield the exact same model-predictions. More generally, I will say that a specific model-outcome *identifies* the model-parameters if and only if a unique combination of parameters predicts that outcome. Note that it is possible that certain outcomes identify the model's parameters, whereas others do not. In practice, model parameters are typically estimated via model-fitting procedures, which aim at finding parameters that minimize the "distance" between the model-outcome and an empirical dataset (e.g., deviance). I will say that an *empirical* dataset identifies the model-parameters, if the minimal distance is obtained for a unique parameter-set. If a model is non-identifiable, different parameter-combinations may yield the minimal distance and hence, the 'true' operative parameters cannot be determined.

[1] Here, I focus mainly on global identifiability, so unless explicitly stated otherwise, the term "identifiability" will imply the global sense.