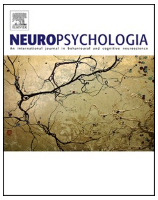
Neuropsychologia 166 (2022) 108114

Contents lists available at ScienceDirect

Neuropsychologia

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



The new hypothesis of everyday amnesia: An effect of criterion placement, not memory

Adva Levi[a], Laura Mickes[b], Yonatan Goshen-Gottstein[a,*]

[a] *Tel-Aviv University, Ramat Aviv, Israel*
[b] *University of Bristol, United Kingdom*

ARTICLE INFO

Keywords
Everyday amnesia
Recognition memory
Bias
Signal detection theory
SDT

Do healthy adults have pockets of anterograde amnesia for information studied several minutes earlier? According to a recent Neuropsychologia article, the answer may be 'yes' (Roediger and Tekin, 2020a; R&T). R&T demonstrated that in a recognition memory test, an old item studied 10 minutes earlier may erroneously be judged as 'new.' The error of not recognizing a studied item—a 'miss'—is not novel. What is novel is the finding that a large proportion of misses, between 16% and 20%, were high-confidence misses (HCMs) (see R&T Table 1). R&T suggest that HCMs are corroborated by anecdotal evidence, as when people encode and recall an event at one time and forget it, completely, soon thereafter. R&T hypothesized that high-confidence misses (HCMs) represent 'everyday amnesia' (EA).

We argue that amnesia is not the driver of HCMs. Instead, HCMs are a consequence of responding and are anticipated by signal-detection theory (SDT). SDT conceptualizes recognition memory as having two components: long-term memory representations for targets and lures[1] (memory) and a decision rule that operates upon these representations (bias). We first show how bias, not memory, accounts for HCMs and then list several concerns regarding the EA hypothesis.

1. Signal-detection theory applied to recognition memory

SDT is premised on the notion that recognition memory is error-prone. Items are sampled from either a continuous distribution of the mnemonic strength of lures or targets (see Fig. 1). For the memory component, participants evaluate the amount of mnemonic evidence available for any test item and use this evidence to judge it as 'old' or 'new.' The response-bias component entails that participants define a magnitude of mnemonic information—by placing a criterion—above which items will be judged 'old' and below which will be judged 'new.' Participants naturally vary in their response bias (i.e., the placement of a criterion).

Misses are incorrect 'new' judgments to targets, and false alarms (FA) are incorrect 'old' judgments to lures. Critically, so long as the lure and the target distributions overlap, these errors are inevitable. The degree of distributional overlap reflects memory strength, whereby the greater the overlap, the weaker the memory.[2]

Thus, SDT conceives both misses and FAs as epiphenomenal to the amount of overlap of the distributions and criterion placement.[3]

To interpret confidence judgments, SDT assumes the placement of

---

; EA, everyday amnesia; HCMs, high-confidence misses; SD, standard deviation; SDT, signal-detection theory; UVSD model, unequal variance signal-detection model.

* Corresponding author. Department of Psychology, Tel-Aviv University, Ramat Aviv 69978, Israel.
E-mail address: goshengott@gmail.com (Y. Goshen-Gottstein).

[1] We know of no evidence that HCMs are restricted to recency items (e.g., Davelaar et al., 2005) and thus assume that they represent a long-term memory phenomenon.

[2] The percent of errors increases with greater overlap and wider lure and target distributions.

[3] Note that a bias interpretation is just as pertinent to recall as it is to recognition, in that recall too is affected by participants' bias in recalling items (Erdelyi et al., 1989). As such, bias provides a coherent interpretation not only of HCMs, but also of the anecdotal evidence described by R&T.

https://doi.org/10.1016/j.neuropsychologia.2021.108114
Received 7 December 2020; Received in revised form 18 February 2021; Accepted 7 December 2021

multiple response criteria (see Fig. 1). In R&T's recognition experiments, participants made old-new judgments followed by confidence ratings. R&T collapsed the ratings to low, medium, and high confidence, yielding six response categories: three for items judged 'old' and three for items judged 'new.' These categories reflect a set of five criteria, C1–C5, ranging from the lowest (most liberal) to the highest (most conservative) along the mnemonic-evidence axis. For example, high confidence 'old' judgments are made when the evidence exceeds the highest criterion, C5. As for the protagonists of this commentary, high-confidence 'new' judgments are made when the evidence does not exceed C1, and constitute HCMs if they were sampled from the target distribution. As we demonstrate, all HCMs effects can be interpreted by the placement of C1—and thus constitute bias effects.

Note that our bias interpretation is impartial to the decision goal that mediate criteria placement (e.g., maximizing proportion correct; see Macmillian and Creelman, 2004, pp. 42–44). Irrespective of the goal, one criterion will inevitably be placed below the others. Items that are sampled from the target distribution and whose strength falls below that criterion, C1, constitute HCMs.

In the following sections, for three findings that might represent part of EA, we ascribe an interpretation entirely dependent on the placement of the C1 criterion. We also report results from Monte Carlo simulations as a feasibility test of the bias account. For all simulations, the critical factor is the proportion of the target distribution that falls to the left of C1. Importantly, to account for HCMs, SDT must provide an unequivocal prediction of the expected *pattern* of HCMs. The precise proportions of HCMs and positions of C1 are not predicted a-priori, and are sample-dependent.

For the simulations, we used the Unequal Variance Signal Detection model (UVSD; see Fig. 1; Egan, 1958; Moran and Goshen-Gottstein, 2015; Wixted, 2007) that we believe provides the best account of recognition memory. In the UVSD model, lures and targets are distributed in Gaussian form, with the target distribution having a larger standard deviation (SD) than the lure-distribution (see Fig. 1). In empirical data, the target SD is often approximately 25% larger (e.g., Mickes et al., 2007; Ratcliff et al., 1992). Unless otherwise noted, for all simulations, the lure mean and SD were set to 0 and 1, respectively, and the target mean and SD were set to 1 and 1.25, respectively (henceforth, 'classic parameters'). We set the criteria equally spaced in steps of 0.5, with the middle criterion set at 0.5 (for further details, see Supplementary Materials A; SMA).

2. The very existence of HCMs

As part of EA, R&T discovered HCMs in healthy adults who fully encoded items just a few minutes before taking the test. Such forgetting made with surprising high confidence led R&T to posit their EA hypothesis. According to SDT, however, HCMs reflect studied target items with strength falling below C1, thus judged 'new' with high confidence (e.g., the red regions in Fig. 1A–C), and their existence is, therefore, predicted.

2.1. Monte Carlo simulations

The mean proportion of HCMs was 15.9%, similar to those reported in Table 1 of R&T (Tekin and Roediger, 2017, Experiment 1; DeSoto and Roediger, 2014, Experiment 1). Furthermore, when the target SD was changed to 1.40, the proportion of HCMS increased to over 21%, similar to experiments reported in Table 1 of R&T (Tekin and Roediger, 2017, Experiment 2). This pattern is expected because the greater variance of the target distribution yields a longer tail, with more items falling below C1. The same increase in HCMs was obtained when but C1 was placed higher (for target SD = 1.25), thus enveloping a higher proportion of targets below it (similar to the examples in Fig. 1C vs. Fig. 1A; see SMA bottom panel). Critically, both interpretations relate to bias, not amnesia.

3. Individual differences in the occurrence of HCMs

R&T suggested that EA may be affected by individual differences, with some participants showing little to no HCMs. According to SDT, as C1 is placed further to the right along the mnemonic strength axis, a higher proportion of HCMs are predicted. As it is placed further to the left, fewer or no targets will fall to its left, yielding a complete absence of HCMs. SDT thus predicts that differences in C1 placement will yield individual differences in the prevalence of HCMs, even when maintaining the same amount of overlap between targets and lures (that is, the same level of memory). We suggest that such differences would reflect differences in criterion placement, not memory.

3.1. Monte Carlo simulations

We simulated 96 participants on 150 test trials, for a total of 7200 lures and 7200 targets (Tekin and Roediger, 2017, Experiment 2, Table 1). Our goal was to show how criterion placement yields some participants with HCMs and some without (i.e., individual differences). An absence of HCMs was defined either in a strict manner—zero HCMs—or more leniently— five-or-fewer HCMs. When placing C1 at -1.25, 6.5% of simulated participants made zero HCMs, with the remaining 93.5% making 1 or more HCMs. When including participants with five-or-fewer HCMs, 41.9% of participants showed a near absence of HCMs, with the remaining 58.1% making five-or-more HCMs. Thus, we demonstrated individual differences in HCMs. However, bias, not memory, mediated these differences. See SMB for more simulations revealing a smaller proportion of participants with HCMs when C1 was set at -1.0 or -1.1.

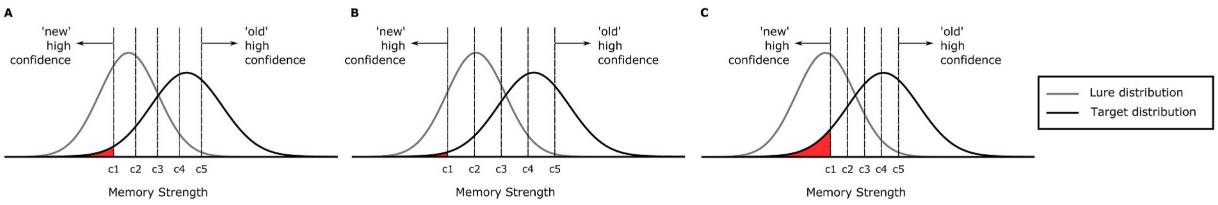


Fig. 1. An illustration of participants' underlying lure and target distributions, as assumed by the unequal variance signal detection (UVSD) model. The left distribution is that of the lures and the right is that of the targets. Five criteria spread across the memory strength axis, representing participants' "cut-offs" for the different confidence judgments. Irrespective of study status (studied, unstudied), items with strength exceeding the highest criterion, C5, are judged 'old' with high confidence. Items with strength between C4 and C5 are judged 'old' with medium confidence, and items with strength between C3 and C4 are judged 'old' with low confidence. Similarly, items judged 'new' received ratings of high, medium, or low confidence for strengths lower than C1, between C1 and C2, and between C2 and C3, respectively. In each panel, the probability of a high-confidence 'new' judgment for targets (high-confidence misses, HCM), is presented as the area under the target distribution and below the criterion C1 (marked red). The difference between the panels, is the criterion shifting, which results in a higher (C) or lower (B) proportion of HCM. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

4. Larger occurrence of HCMs for better memory

According to R&T's EA hypothesis, an increase in HCMs may be observed when memory performance is better. Roediger and Tekin (2020b) compared two studies and noted a higher proportion of HCMs in the study with better memory. According to SDT models, changes in the placement of C1 are made independently of the amount of overlap between the distributions (cf., Stretch and Wixted, 1998). Thus, the incidence of HCMs is dependent on the strength of the memories only if a manipulation of memory strength happens to incidentally also affect bias.

4.1. Monte Carlo simulations

We compared the 15.9% HCMs obtained in our simulation using the classic parameters, with C1 = −1.0, to a larger target mean of 1.5—corresponding to less of an overlap of the distributions and representing superior memory. When C1 was set to a low value of −1.75, only 2.2% HCMs were found. Thus, better memory was associated with fewer HCMs. This was reversed when, keeping the target mean at 1.5, we set C1 to −0.25, yielding 38.1% HCMs, thereby demonstrating more HCMs with better memory (see SMC). Once again, a bias interpretation accounts for the different patterns reported by R&T.

5. Strength of SDT and weakness of the EA interpretation

SDT is a well-defined, mathematically specified theory. While flexible, it can be tested and is falsifiable. One such test is the strong commitment to the existence of errors, be they high-confidence or low-confidence FAs and misses. The absence of such errors would undeniably provide a refutation of SDT (assuming, of course, sufficient power). SDT has been put to the test time after time, with the most recent example "accepted for publication"—in arguably the most theoretically rigorous journal of scientific psychology—as these words are being written. Kellen et al. (2021) tested five fundamental assumptions of SDT recognition models, with each assumption passing its respective test.

We have several concerns with the interpretation of HCMs as EA, all of which are addressed by (or not relevant for) an SDT interpretation. First, according to EA, what is the latent variable upon which participants make confidence judgments? Second, is the latent variable continuous? If so, what are the decision rules for mapping the continuous latent variable onto discrete categories of confidence? Third, can errors occur under these decision rules? Fourth, if they can, should they not be interpreted in terms of bias rather than amnesia? Fifth, the notion of amnesia implies that items are either not available or inaccessible. Why not attribute the very same set of findings to decision rules? Finally, should criteria not be articulated and shown to be fulfilled as a prerequisite for the postulation of new forms of amnesia (e.g., proposal raised as a prerequisite for the postulation of distinct memory systems; Roediger et al., 1990)?

Occam's razor requires that explanations of unknown phenomena be sought first in terms of known quantities. SDT is undeniably a known quantity and is regarded as one of the most successful (for reviews, see Kellen and Klauer, 2018; Wixted, 2020) and universally accepted (Estes, 2002) theoretical frameworks. It is these considerations that lead us to embrace a bias interpretation afforded by SDT, as the most viable account of HCMs.

Acknowledgements

The Commentary is part of A.L.'s doctoral dissertation. This work was supported by the Israel Science Foundation [Grant number 1163–20].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2021.108114.

References

Davelaar, E.J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H.J., Usher, M., 2005. The demise of short-term memory revisited: empirical and computational investigations of recency effects. Psychol. Rev. 112, 3–42.

DeSoto, K.A., Roediger III, H.L., 2014. Positive and negative correlations between confidence and accuracy for the same events in recognition of categorized lists. Psychol. Sci. 25, 781–788.

Egan, J.P., 1958. *Recognition Memory and the Operating Characteristic* (Tech. Note AFCRC-TN-58–51). Indiana University, Hearing and Communication Laboratory, Bloomington.

Erdelyi, M.H., Finks, J., Feigin-Pfau, M.B., 1989. The effect of response bias on recall performance, with some observations on processing bias. J. Exp. Psychol. Gen. 118, 245–254.

Estes, W.K., 2002. Traps in the route to models of memory and decision. Psychonomic Bull. Rev. 9, 3–25.

Kellen, D., Klauer, K.C., 2018. Elementary signal detection and threshold theory. In: Wixted, J.T. (Ed.), Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience. Wiley, pp. 1–39.

Kellen, D., Winiger, S., Dunn, J.C., Singmann, H., 2021. Testing the foundations of signal detection theory in recognition memory. Psychol. Rev.

Macmillan, N.A., Creelman, C.D., 2004. Detection Theory: A User's Guide. Lawrence Erlbaum, Mahwah, NJ.

Mickes, L., Wixted, J.T., Wais, P.E., 2007. A direct test of the unequal-variance signal detection model of recognition memory. Psychonomic Bull. Rev. 14, 858–865.

Moran, R., Goshen-Gottstein, Y., 2015. Old processes, new perspectives: familiarity is correlated with (not independent of) recollection and is more (not equally) variable for targets than for lures. Cognit. Psychol. 79, 40–67.

Ratcliff, R., Sheu, C., Gronlund, S.D., 1992. Testing global memory models using ROC curves. Psychol. Rev. 99 (3), 518–535.

Roediger III, H.L., Rajaram, S., Srivivas, K., 1990. Specifying criteria for postulating memory systems. In: Diamond, A. (Ed.), The Development and Neural Bases of Higher Cognitive Functions. New York Academy of Sciences Press, New York, pp. 572–595.

Roediger, H.L., Tekin, E., 2020a. Recognition memory: tulving's contributions and some new findings. Neuropsychologia 139, 107350.

Roediger, H.L., Tekin, E., 2020b. November 19-21). *Everyday amnesia: High confidence misses in recognition memory* [Conference presentation]. Psychonomic Society 2020 Annual Virtual Meeting.

Stretch, V., Wixted, J.T., 1998. Decision rules for recognition memory confidence judgments. J. Exp. Psychol. Learn. Mem. Cognit. 24, 1397–1410.

Tekin, E., Roediger, H.L., 2017. The range of confidence scales does not affect the relationship between confidence and accuracy in recognition memory. Cognitive Research: Principles and Implications 2, 49.

Wixted, J.T., 2007. Dual-process and signal detection theory of recognition memory. Psychol. Rev. 114, 152–176.

Wixted, J.T., 2020. The forgotten history of signal detection theory. J. Exp. Psychol. Learn. Mem. Cognit. 46, 201–233.