



Age differences in diffusion model parameters: a meta-analysis

Maximilian Theisen¹ · Veronika Lerche¹ · Mischa von Krause¹ · Andreas Voss¹

Received: 5 July 2019 / Accepted: 4 June 2020 / Published online: 13 June 2020
© The Author(s) 2020, corrected publication 2021

Abstract

Older adults typically show slower response times in basic cognitive tasks than younger adults. A diffusion model analysis allows the clarification of why older adults react more slowly by estimating parameters that map distinct cognitive components of decision making. The main components of the diffusion model are the speed of information uptake (drift rate), the degree of conservatism regarding the decision criterion (boundary separation), and the time taken up by non-decisional processes (i.e., encoding and motoric response execution; non-decision time). While the literature shows consistent results regarding higher boundary separation and longer non-decision time for older adults, results are more complex when it comes to age differences in drift rates. We conducted a multi-level meta-analysis to identify possible sources of this variance. As possible moderators, we included task difficulty and task type. We found that age differences in drift rate are moderated both by task type and task difficulty. Older adults were inferior in drift rate in perceptual and memory tasks, but information accumulation was even increased in lexical decision tasks for the older participants. Additionally, in perceptual and lexical decision tasks, older individuals benefitted from high task difficulty. In the memory tasks, task difficulty did not moderate the negative impact of age on drift. The finding of higher boundary separation and longer non-decision time in older than younger adults generalized over task type and task difficulty. The results of our meta-analysis are consistent with recent findings of a more pronounced age-related decline in memory than in vocabulary performance.

Introduction

It is a common finding from the literature on cognitive aging that older people show larger response times (RTs) in basic cognitive tasks than younger adults (Jensen, 2006). In the last decades, the mechanisms underlying this age-related slowing have become a subject of debate. On the one hand, the higher RTs of the older adults might be the result of a general decline in cognitive processing speed due to increased neural noise (Salthouse, 1996). On the other hand, however, it is also possible that the slow responses are based on encoding problems (e.g., due to impaired vision), reduced motoric speed, or more cautious response criteria. Such different accounts can be differentiated by means of diffusion model analyses (Voss, Nagler, & Lerche, 2013). The diffusion model (Ratcliff, 1978) is a stochastic model used to analyze response time distributions and error rates

in binary decision tasks. It thus utilizes a more complete representation of decision outcomes than just mean RTs. The model aims to disentangle three main components of the decision process: the speed of information uptake (*drift rate*), the degree of conservatism regarding the decision criterion (i.e., speed-accuracy trade-offs; *boundary separation*) and the time taken up by non-decisional processes such as encoding and motoric response execution (*non-decision time*).

Several diffusion model studies have challenged the view that age differences in RT are indicative of a general decline in cognitive speed (e.g., Spaniol, Madden, & Voss, 2006). Quite often, age differences in RT were not due to differences in the mean speed of information uptake, but due to the fact that older people tended to be more cautious (i.e., they favored accurate over fast responses) and that they took longer in terms of the non-decisional components of the response time (e.g., Ratcliff, Thapar, & McKoon, 2001). However, in some studies, older people additionally showed a lower speed of information uptake (Voskuilen, Ratcliff, & McKoon, 2018), consistent with the notion that processing speed declines with age.

✉ Maximilian Theisen
maximilian.theisen@psychologie.uni-heidelberg.de

¹ Psychologisches Institut, Ruprecht-Karls-Universität Heidelberg, Hauptstrasse 47-51, 69117 Heidelberg, Germany

So far, to our knowledge no attempt has been made to bring together the inconsistent results regarding drift rates in a quantitative way. It is an open question whether the discrepancies are simply due to random sample differences or can be explained by specific study attributes. As Dully, McGovern, and O'Connell (2018) note in their literature review, there are “task-specific differences in evidence accumulation rates” (p. 3). However, these task-specific differences have not yet been examined quantitatively.

In this paper, we present the results of a meta-analysis regarding age differences in diffusion model parameters. The focus is on drift rates because of the variability in findings for this parameter. We were interested in whether characteristics of the task (specifically, content and difficulty of task) might explain the inconsistent findings in the literature. We also analyzed the parameters boundary separation and non-decision time. In terms of these parameters, we expected that age differences generalize across tasks. In the next chapter, we briefly introduce the diffusion model (for further introductory information, see e.g., Ratcliff & McKoon, 2008; Voss et al., 2013; Wagenmakers, 2009).

Introduction to diffusion modeling

The diffusion model is a mathematical model that can be applied to examine the processes underlying RT tasks with two response options. It has most frequently been used with three main task types (Voss et al., 2013). The first group of tasks comprise *memory tasks*. Here, participants usually have to decide whether a stimulus has been presented to them before or not (recognition memory tasks, e.g., Spaniol et al., 2006; Yap, Sibley, Balota, Ratcliff, & Rueckl, 2015). Second, there are *perceptual tasks* in which participants have to discriminate, for example, between two levels of brightness (bright vs. dark, e.g., Ratcliff, 2002; Ratcliff, Thapar, & McKoon, 2003), between two different letters (e.g., F vs. Q, Thapar, Ratcliff, & McKoon, 2003) or between two different quantities of stimuli (small vs. large, e.g., Ratcliff, Thompson, & McKoon, 2015; Ratcliff & Van Dongen, 2009). The third category of task types includes *lexical decision tasks*. In these tasks, participants have to assess whether a presented letter string is a word or not (e.g., Ratcliff, Gomez, & McKoon, 2004; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008).

For these three categories of tasks, it is assumed that the four central assumptions of the diffusion model are met: (1) Information about the two response options is accumulated continuously, (2) decisions are based on single-stage processing, (3) parameters are constant over time, and (4) the tasks are fast response time tasks with mean RTs below 1.5 s. Note, however, that this latter criterion has recently been questioned. Studies demonstrated that also for RT tasks

that take up to several seconds per trial, the diffusion model fits well and provides valid parameter estimates (Aschenbrenner, Balota, Gordon, Ratcliff, & Morris, 2016; Lerche, Christmann, & Voss, 2018; Lerche & Voss, 2017). In fact, for such slower tasks, the standard diffusion model that is based on random Gaussian noise might even fit better than for very fast response time tasks (Voss, Lerche, Mertens, & Voss, 2019).

Four main parameters affect the position and shape of response time distributions in the diffusion model framework. These parameters are also visualized in Fig. 1. First, there is the distance between the two boundaries that are associated with correct (upper boundary) and error responses (lower boundary) in the example figure. This *boundary separation* (a) defines the quantity of information that needs to be accumulated before a decision is made. Under accuracy (speed) instructions, participants typically adopt more distant (more close) boundaries (e.g., Ratcliff & Rouder, 1998; Voss, Rothermund, & Voss, 2004). Second, there is the speed of information accumulation, the so-called *drift rate* (ν). Drift rate is higher in easier compared to more difficult tasks (e.g., Arnold, Bröder, & Bayen, 2015; Lerche & Voss, 2017) and drift is also positively related to cognitive abilities (e.g., Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007; Schubert, Hagemann, Voss, Schankin, & Bergmann, 2015).

Third, the *starting point* (z) of the accumulation process is modelled. In many tasks, decision processes start from the center between the two boundaries. However, if one of the two response options has a higher expected value (e.g., the

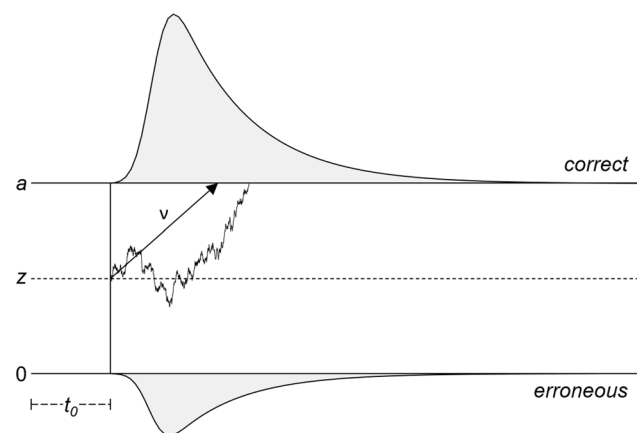


Fig. 1 Diffusion model with its four main parameters. The boundaries are associated with correct and erroneous responses here. One exemplary trial is illustrated. In this trial, the accumulation process starts at starting point z , which is here right in the center between the two boundaries (0 and a). The process moves with speed ν toward the upper boundary. To this straight process adds random Gaussian noise. For convenience, parameter t_0 is illustrated at the left from the decision process. Note that it also includes processes succeeding the decisional process (the motoric response)

response is correct in more trials or higher reward is associated with this response), participants shift the starting point towards the favored option (e.g., Leite & Ratcliff, 2011; Voss et al., 2004).

Finally, *non-decision time* (t_0) subsumes the total duration of all the non-decisional processes (e.g., encoding of information and motor response). Moreover, inter-trial variabilities are often included in the model. However, these variability parameters (in particular, inter-trial variability of drift rate and starting point) often cannot be estimated very reliably (Boehm et al., 2018; Lerche & Voss, 2016; van Ravenzwaaij, Donkin, & Vandekerckhove, 2017) and thus are not very useful to assess inter-individual differences in decision making.

Methods

As the main focus of this meta-analysis is on examining inconsistent findings concerning age differences in drift rate, the literature search concentrated on studies comparing mean drift rates between two age groups. For these studies, we additionally coded effect sizes for boundary separation and non-decision time. Below, we report our procedure in detail.

Inclusion criteria and literature search

For our literature search, we used the following two inclusion criteria:

1. All studies had to refer to the original publication introducing the diffusion model in psychology (Ratcliff, 1978) and they had to report results from a diffusion model analysis. Articles applying the EZ-diffusion model (Wagenmakers, van der Maas, & Grasman, 2007) were included. However, we excluded all studies in which parameter estimation was based on fitting the Ex-Gaussian or the shifted Wald distributions due to concerns about the interpretability of their parameters (Matzke & Wagenmakers, 2009).
2. The second required inclusion criterion regards an obligatory comparison between younger adults (college age) and healthy older adults (youngest participant older than 55 or mean age > 60). We excluded studies reporting continuous age analyses if no categorical age data could be extracted from the reported results (e.g., the relation between age and the corresponding parameter of the diffusion model was only provided as a correlation, without raw data being available). We included studies reporting results from more than two age groups if college-aged adults and older adults were among these groups. In case

of two higher age groups, we used only the younger one of them to enhance comparability between studies.

We used Google Scholar's search engine to collect studies, as it allows to combine a descendant approach with the use of specific keywords (for the comparability of Google Scholar with established scientific databases, see Anders & Evans, 2010; Gehanno, Rollin, & Darmoni, 2013; Shultz, 2007). In a first step, we identified all the papers citing Ratcliff's seminal work on the diffusion model (1978) ($k = 3341$). The next step consisted in searching these studies using age-related terms,¹ resulting in $k = 561$ publications. The search was finished on January 16, 2019.

We conducted a full-text scan of these papers searching for studies that fulfilled the inclusion criteria, resulting in $k = 48$ articles. After removal of duplicates, $k = 46$ articles remained. Several articles did not report sufficient data to calculate effect sizes on drift rate, resulting in a final dataset of 21 papers. Some papers reported data from more than one sample (e.g., if more than one experiment conducted on different participants is reported in the same publication) and/or more than one effect size per sample (e.g., if different tasks were reported for the same participants). We retrieved effect sizes from 25 samples. For boundary separation and non-decision time, we had to exclude one sample, respectively, as the reported data was not sufficient. In total, we retrieved 146 effect sizes for drift rate, 47 effect sizes for boundary separation, and 40 effect sizes for non-decision time.

Calculation of effect sizes

As effect size measure we used Hedges' g (Hedges, 1981). We computed effect sizes using the *compute.es* package (version 0.2-4; Del Re, 2013) of the *R* open-source software environment (version 3.5.1; R Core Team, 2018). If a paper did not report means or standard deviations, we used inferential statistics to determine effect sizes. Positive effect sizes indicate higher values for higher age.

Coding of moderator variables

For each study, we coded the *type of task* using the categories described in Voss, Nagler, and Lerche (2013). Following this classification, most binary decision tasks analyzed with the diffusion model are either perceptual, lexical decision or memory tasks. Some experimental tasks did not fit

¹ search string: ("age differences" OR "old adults" OR "old participants" OR "older adult" OR "older adults" OR "older participants" OR "higher age" OR "older group" OR "old group" OR "age-related" OR "effects of aging" OR "effects of age" OR "aging effects" OR "age effects").

Table 1 Samples included in the final dataset of the meta-analysis

Articles	<i>n</i>	<i>n</i> young	<i>n</i> old	Age range young	Age range old	Mean age young	Mean age old
Allen, Lien, Ruthruff, and Voss (2014)	21	11	10	18–24	64–80	21.7	71.8
Ball and Aschenbrenner (2018)	125	67	58	18–21	60–90	18.9	75.0
Dirk et al. (2017)	40	20	20	18–36	64–75	25.7	68.1
Huff and Aschenbrenner (2018)	163	85	78			21	74.6
Kapucu (2010)	56	30	26			19.8	71.9
Kordella (2009)							
Experiment 2	41	22	19	18–24	61–74	20.2	68.9
Experiment 3	38	22	16	18–25	60–74	20.1	68.3
Kühn et al. (2011)	39	24	15	20–31	65–80	25.2	70.2
McKoon and Ratcliff (2012)	78	39	39	18–25	60–74	20.6	68.4
McKoon and Ratcliff (2013)	67	30	37		60–74	20.8	69.7
Ratcliff (2008)	38	19	19		60–75	20.8	69.2
Ratcliff, Thapar, Gomez, and McKoon (2004)							
Experiment 1	98	54	44			19.8	68.5
Experiment 2	94	54	40			20.2	67.2
Ratcliff, Thapar, and McKoon (2004)	80	39	41		60–74	19.6	70
Ratcliff, Thapar, and McKoon (2006)	20	10	10		60–74		
Ratcliff, Thapar, and McKoon (2010)	88	45	43	18–25	60–74	20.8	68.6
Ratcliff, Thapar, and McKoon (2011)	91	46	45		60–74	20.4	68.3
Spaniol, Voss, and Grady (2008)							
Experiment 1	43	22	21	19–28	60–75	22.5	67.5
Experiment 2	47	24	23	18–32	61–85	22.3	71.8
Spaniol, Voss, Bowen, and Grady (2011)	53	26	27	18–32	61–85	23.0	71.5
Thapar, Ratcliff, and McKoon (2003)	78	40	38		60–75	19.8	69.1
Voskuilen et al. (2018)	23	11	12		60–80		

this classification scheme. We omitted the according effect sizes from the analyses (11 effect sizes for drift rate, 2 effect sizes for boundary separation and 4 effect sizes for non-decision time). See Table 1 for the articles included in this final dataset.

A second moderator variable in our analyses was *task difficulty*. We used drift rate as measure of task difficulty as the literature suggests that more difficult tasks go along with lower drift rates (e.g., Arnold, Bröder, & Bayen, 2015; Voss et al., 2004). In several studies, task difficulty varied between conditions. Here we computed a mean drift rate across the different difficulty levels and age groups (weighting by the number of participants per group). Next, we *z*-transformed and inverted the mean drift so that higher values of the variable indicate enhanced task difficulty.

Statistical analyses

As several effect sizes are based on the same samples, we assumed effect sizes to be dependent. We accounted for this dependent structure by conducting multilevel meta-analyses using the *metafor* package (version 2.0-0; Viechtbauer, 2010) in *R*. We specified the levels as effect size

nested in sample with task type as an inner grouping factor. This means that effects stemming from different samples are assumed to be independent, while effects of the same task type within a sample share correlated random effects. The variance structure of the inner factor was assumed to be a heteroscedastic compound symmetric structure.²

We used the maximum likelihood estimation procedure included in the function `rma.mv()` and analyzed the three outcome variables in independent sets of analyses (i.e., drift rate, boundary separation, and non-decision time). In a first step, we ran multilevel meta-analyses without any moderators (Model 1). Then, in a second step, we included task type and task difficulty as moderators (Model 2). As we were also interested in a possible interaction between task type and task difficulty, we further added the interactions in a third step (Model 3).

The validity of meta-analytical models can suffer because of influential outliers. To date, the development of tools for outlier and influence diagnostics for multilevel

² As recommended by Wolfgang Viechtbauer (Personal communication, April 2018).

Table 2 Number of available effect sizes for each diffusion model parameter depending on the task type

Parameter	Perceptual tasks	Lexical decision tasks	Memory tasks
Drift rate	30	16	89
Boundary separation	16	6	23
Non-decision time	14	6	16

meta-analyses is still in progress (Viechtbauer & Cheung, 2010). We followed the procedure of Habel and Schultz (2015), removing any influential outliers, defined by effect sizes with both hat values greater than two times the average hat value and standardized residual values exceeding 3.29.

We tested for publication bias using Egger's regression test (Egger, Smith, Schneider, & Minder, 1997; Sterne & Egger, 2005) by modifying Model 1 to include the variance of the effect size as moderator (Moreno et al., 2009). If the intercept of this model significantly deviates from zero, the relationship between variance and effect size can be assumed to be asymmetrical, indicating a bias. Because of the low power of this test for publication bias, we set the alpha-level to $\alpha = 0.10$ (Egger et al., 1997).

Furthermore, we assessed heterogeneity among effect sizes using Cochran's Q statistic and the I^2 statistic. Large Q values indicate that differences among effect sizes can be attributed to differences among the true effects and do not solely result from sampling errors. If the Q test is significant, the integrated effect size is not an estimator of the true effect but rather an estimator of the mean of the distribution of different true effect sizes (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Results

Study characteristics

The included studies stem from the period of 2003–2018. In total, we analyzed the data of 1503 participants ($M = 62.63$ per sample, $SD = 34.90$). The mean age of the young groups was 21.15 ($SD = 1.75$), the mean age of the older groups was 69.77 ($SD = 2.17$). Table 2 shows the distribution of task types over the three diffusion model parameters (see Table S1 in the Supplementary Material for a detailed description of the respective task and condition for each included effect size).

Diagnostics

For drift rate, there were two cases with standard residual values greater than 3.29. However, their hat values did not

exceed 2 and, therefore, we did not omit them from the analyses. For boundary separation and non-decision time, we found no outliers. The intercepts of Egger's regression models indicated publication bias for all three diffusion model parameters (drift rate: $\beta_0 = 1.049$, $p < 0.001$; non-decision time: $\beta_0 = 0.988$, $p < 0.001$; boundary separation: $\beta_0 = -0.637$, $p = 0.063$).

Meta-analysis

Drift rate

The meta-analytical model with task type and task difficulty as moderators (Model 2; $AICc = 376.2$) had a better fit than the model without moderators (Model 1; $AICc = 379.1$, $p = 0.022$). Including the interaction between task type and difficulty improved the model fit even further (Model 3; $AICc = 374.1$, $p = 0.034$). Thus, our final meta-analytical model contained task type, task difficulty, and their interaction as moderator variables (see Figure S1 in the Supplementary Material for a forest plot of the final model). For the final model, the Q test was highly significant, $Q(129) = 1016.331$, $p < 0.001$. The estimated standard deviations of true effects per task type were $\tau = 0.848$ (perceptual tasks), $\tau = 1.153$ (lexical decision tasks), and $\tau = 0.549$ (memory tasks). The I^2 values for the three levels of task type were 91.61% (perceptual tasks), 95.28% (lexical decision tasks), and 82.07% (memory tasks).³

The mean effect sizes per task type were $g = -0.608$, 95% CI $[-1.032, -0.184]$, $p = 0.005$ for perceptual tasks, $g = 0.620$, 95% CI $[0.037, 1.203]$, $p = 0.037$ for lexical decision tasks and $g = -0.326$, 95% CI $[-0.587, -0.065]$, $p = 0.014$ for memory tasks (see Fig. 2 for a graphical representation). This indicates reduced drift rates for older adults in perceptual and memory tasks but increased drift rates of older adults for lexical decision tasks. Furthermore, in more difficult tasks older adults performed relatively better compared to younger adults ($\beta = 0.181$, $p = 0.010$).

To examine the task type by task difficulty interaction, we additionally conducted separate analyses for each type of task, with and without task difficulty as moderator. We then compared the fit of the model with and without difficulty to test if this moderator explains variance within a task type (Table 3). For perceptual and lexical decision tasks, the model with difficulty as moderator performed significantly better than the model without moderator. More specifically, task difficulty significantly predicted effect sizes for

³ To compute the I^2 on task type level we used the approach for multivariate models as described in Viechtbauer (2018, December 8). To compute the I^2 for boundary separation and non-decision time, we used Higgins and Thompson's (2002) formula.

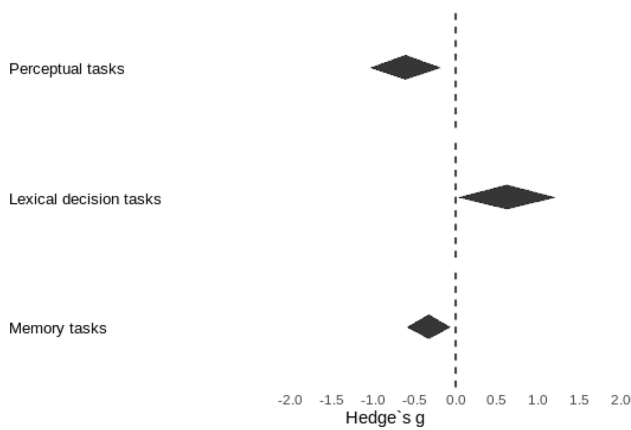


Fig. 2 Mean age effects in drift rate for each of the task types analyzed. 95% confidence intervals indicated by the width of the polygons

Table 3 Drift rate: comparisons between models with and without task difficulty as moderator for each task type subset

Statistic	Perceptual tasks	Lexical decision tasks	Memory tasks
AICc _{with task difficulty}	69.710	63.887	244.189
AICc _{without task difficulty}	75.465	65.395	242.100
<i>p</i>	.004	.032	.817

perceptual ($\beta = 0.203, p = 0.004$) and lexical decision tasks ($\beta = 0.719, p = 0.022$): Older adults profited from high task difficulty. For memory tasks, on the other hand, task difficulty did not predict effect sizes, $\beta = 0.016, p = 0.816$. In the supplementary materials, we provide a full forest plot showing all drift rate effect sizes analyzed.

Boundary separation

The meta-analytical model without any moderators (AICc = 143.8) showed a better fit to the data than the model with task type and task difficulty as moderators (AICc = 149.1, $p = 0.372$). Therefore, we kept the model without any moderators. The mean effect size of age on boundary separation was $g = 0.731, 95\% \text{ CI } [0.472, 0.989], p < 0.001$. Results indicate that older adults generally adopt higher boundary separations (i.e., a more conservative response criterion) than young adults. The *Q* test was highly significant, $Q(44) = 669.203, p < 0.001; I^2$ for the whole model was 93.13%.

Non-decision time

The meta-analytical model without any moderators (AICc = 89.9) showed a better fit than the model with task

type and task difficulty (AICc = 96.1, $p = 0.379$). Therefore, we kept the model without any moderators. The mean effect size of age on non-decision time was $g = 1.673, 95\% \text{ CI } [1.404, 1.942], p < 0.001$. Our results suggest that older people show a longer non-decision time than younger people. The *Q* test was highly significant, $Q(37) = 388.946, p < 0.001; I^2$ for the whole model was 90.487%.

Discussion

In the last decades, the diffusion model (Ratcliff, 1978) has become a popular approach for the analysis of age differences in response time tasks. The findings from the diffusion model analyses have challenged the view that cognitive processing speed generally declines with age. Rather, the studies revealed a more complex picture, which we wanted to examine further in our meta-analysis. Most importantly, we were interested in the drift rate, which is a measure of speed of information accumulation that is closely related to intelligence (e.g., Ratcliff, Thapar, & McKoon, 2011; Schmitz & Wilhelm, 2016). Regarding age effects on drift, previous studies provided inconsistent findings. Whereas some studies report reduced drift rates for older adults (e.g., Thapar et al., 2003), other studies do not find differences in this model parameter (e.g., Ratcliff et al., 2001), or even higher drift rates for older adults (e.g., Ratcliff, Thapar, & McKoon, 2010). With the present meta-analysis, we aim to identify reasons for this heterogeneity. To this aim, we assembled studies that report drift rate differences between a young (college age) and an old age group (> 55 years). Then, we examined the influence of task difficulty and task type (perceptual, lexical decision, and memory) on age effects in diffusion model parameters. Thus, we could uncover possible important moderators that might explain (part of) the inconsistent findings in the literature.

Boundary separation and non-decision time

Results provided two most clear-cut findings: First, older adults are slower than young adults in non-decisional processes (such as encoding of information and motoric response execution). The corresponding effect size was large ($g = 1.673$). Second, older adults generally use more conservative response criteria (i.e., larger boundary separations) than young adults. Even if the effect size is smaller than for non-decision time, it is still substantial ($g = 0.731$). Thus, older individuals are more cautious in their decisions. These effects did not depend on either task type or task difficulty. Note that both boundary separation and non-decision time influence RT (e.g., Ratcliff & McKoon, 2008). Thus, the common finding of higher RTs of older adults seems to be highly attributable to these two parameters.

Drift rate

Whereas age differences in boundary separation and non-decision time generalized across different task types and difficulties, we found moderator effects for speed of information accumulation (drift rate). In perceptual tasks and memory tasks, older adults had lower drift rates than younger adults. However, the older groups were superior in speed of information accumulation compared to their younger counterparts in lexical decision tasks. Furthermore, task difficulty also influenced age effects on drift: In terms of perceptual and lexical decision tasks, older participants profited from more difficult tasks. In the memory tasks, task difficulty did not moderate the effect of age on drift.

Thus, our study shows that the pattern of results is clearly more complex for drift rate than for boundary separation and non-decision time and that it seems to be important to consider the specific cognitive processes required by different experimental paradigms. In line with this finding are the results from a recent diffusion model study that is based on a set of 18 different RT tasks (Lerche et al., 2020). The study revealed domain-specific drift factors (numeric, verbal, figural) that were further related to the respective components of an intelligence test. Thus, speed of information accumulation seems to be dependent on the task content. Furthermore, also neurophysiological studies found that aging effects depend on the task (Dully et al., 2018).

Our meta-analysis suggests that older adults outperform young adults in lexical decision tasks, whereas they perform worse in memory tasks. This is in line with the findings from studies that date back to the 1920s and 30s (e.g., Conrad, Jones, & Hsiao, 1933; Foster & Taylor, 1920; Willoughby, 1929). The results of these studies suggest that age differences are more pronounced in measures of memory than vocabulary. Also, recent studies generally confirm this observation. For example, Salthouse (2004), aggregating across several studies from 1998 till 2003, reports a substantial, linear age decline in performance in a memory test. On the other hand, performance improved with age in a vocabulary test, at least until about the mid-50s. After that, it remained stable or somewhat declined (confer also Spaniol et al., 2006). Our meta-analysis showed that such task-specificities are captured in the drift rate of the diffusion model. Furthermore, our analysis revealed that not only the type of task, but also task difficulty needs to be considered. Older adults profited from the more difficult tasks.

Limitations and future research

Even if the overall number of effect sizes for drift rate used for the meta-analysis was substantial ($N = 135$), analyses of the moderator task type were based on smaller case

numbers. Here, the distribution was not balanced with clearly fewer lexical decision effect sizes ($n = 16$) than effect sizes from perceptual ($n = 30$) or memory tasks ($n = 89$). In future research, one might try to replicate our findings in large-scale studies that are explicitly designed to measure the influence of task type (and difficulty) on age differences. Further note that despite consideration of two moderator variables, there was still substantial unexplained variance in our meta-analysis. Accordingly, in future studies, one might try to identify further possible moderators.

The focus of our meta-analysis was on drift rate because findings in the literature seemed to be more inconsistent for this parameter. Therefore, our search strategy was based on finding all studies that report age differences in drift rate. With this strategy we do not identify studies that report age differences only in boundary separation or non-decision time, but not in drift rate. Accordingly, the superiority of the model without moderators might also be partly attributable to the small cell numbers (between 6 and 23 for the different task types). Thus, if one would like to examine moderator influences for non-decision time and boundary separation in more detail, separate meta-analyses should be conducted.

Finally, it would be highly interesting to examine age effects more systematically also for other sequential sampling models, e.g., the popular linear ballistic accumulator model (LBA; Brown & Heathcote, 2008). So far, the literature on age effects in LBA model parameters is more limited than the respective diffusion model literature. The previous LBA findings seem to be generally in line with the results from our meta-analysis: In comparison with younger adults, older adults have been found to have higher threshold separations (Forstmann et al., 2011; Garton, Reynolds, Hinder, & Heathcote, 2019), and longer non-decision times (Ben-David, Eidels, & Donkin, 2014; Garton et al., 2019), whereas the results for drift rate are less clear-cut. Further, previous research suggests that the diffusion model parameters and the LBA model parameters have very similar meanings (Donkin, Brown, Heathcote, & Wagenmakers, 2011). However, to note, in a recent multi-lab project one systematic difference between the two models emerged (Dutilh et al., 2019). More specifically, for instructions that focused either on accuracy or speed the teams that used the diffusion model often found an effect in non-decision time (in addition to an effect in threshold separation), whereas the LBA teams often detected an effect in drift rate. The reasons for this pattern of results will need to be investigated further in future research (for a recent discussion of this topic, see Evans, 2020; Lerche & Voss, 2018). Based on these varying findings, we hypothesize that somehow different age effects might emerge if older and younger adults are compared based on different sequential sampling models. For example, effect sizes for age effects in non-decision time might be larger for the diffusion model than for the LBA.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00426-020-01371-8>) contains supplementary material, which is available to authorized users.

Funding Open access funding enabled and organized by Projekt DEAL. This research was supported by grants from the German Research Foundation to Andreas Voss (Grant No. VO1288/2-2) and to the Graduate School SMiP (GRK 2277; Statistical Modeling in Psychology).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allen, P. A., Lien, M.-C., Ruthruff, E., & Voss, A. (2014). Multitasking and aging: Do older adults benefit from performing a highly practiced task? *Experimental Aging Research*, *40*(3), 280–307. <https://doi.org/10.1007/s00426-014-0608-y>.
- Anders, M. E., & Evans, D. P. (2010). Comparison of PubMed and google scholar literature searches. *Respiratory Care*, *55*(5), 578–583.
- Arnold, N. R., Bröder, A., & Bayen, U. J. (2015). Empirical validation of the diffusion model for recognition memory and a comparison of parameter-estimation methods. *Psychological Research Psychologische Forschung*, *79*(5), 882–898. <https://doi.org/10.1007/s00426-014-0608-y>.
- Aschenbrenner, A. J., Balota, D. A., Gordon, B. A., Ratcliff, R., & Morris, J. C. (2016). A diffusion model analysis of episodic recognition in preclinical individuals with a family history for Alzheimer's disease: The adult children study. *Neuropsychology*, *30*(2), 225–238. <https://doi.org/10.1037/neu0000222>.
- Ball, B. H., & Aschenbrenner, A. J. (2018). The importance of age-related differences in prospective memory: Evidence from diffusion model analyses. *Psychonomic Bulletin & Review*, *25*(3), 1114–1122. <https://doi.org/10.3758/s13423-017-1318-4>.
- Ben-David, B. M., Eidels, A., & Donkin, C. (2014). Effects of aging and distractors on detection of redundant visual targets and capacity: Do older adults integrate visual targets differently than younger adults? *PLoS ONE*, *9*(12), 1–29. <https://doi.org/10.1371/journal.pone.0113551>.
- Boehm, U., Annis, J., Frank, M. J., Hawkins, G. E., Heathcote, A., Kellen, D., et al. (2018). Estimating across-trial variability parameters of the Diffusion Decision Model: Expert advice and recommendations. *Journal of Mathematical Psychology*, *87*, 46–75. <https://doi.org/10.1016/j.jmp.2018.09.004>.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.
- Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178. <https://doi.org/10.1016/j.cogpsych.2007.12.002>.
- Conrad, H. S., Jones, H. E., & Hsiao, H. H. (1933). Sex differences in mental growth and decline. *Journal of Educational Psychology*, *24*(3), 161–169. <https://doi.org/10.1037/h0073913>.
- Del Re, A. C. (2013). compute.es: Compute Effect Sizes. R package version 0.2-2. Retrieved from <https://cran.r-project.org/web/packages/compute.es>
- Dirk, J., Kratzsch, G. K., Prindle, J. P., Kröhne, U., Goldhammer, F., & Schmiedek, F. (2017). Paper-based assessment of the effects of aging on response time: A diffusion model analysis. *Journal of Intelligence*, *5*(12), 1. <https://doi.org/10.3390/jintelligence5020012>.
- Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E.-J. (2011). Diffusion versus linear ballistic accumulation: Different models but the same conclusions about psychological processes? *Psychonomic Bulletin & Review*, *18*(1), 61–69. <https://doi.org/10.3758/s13423-010-0022-4>.
- Dully, J., McGovern, D. P., & O'Connell, R. G. (2018). The impact of natural aging on computational and neural indices of perceptual decision making: A review. *Behavioural Brain Research*, *355*, 48–55. <https://doi.org/10.1016/j.bbr.2018.02.001>.
- Dutilh, G., Annis, J., Brown, S. D., Cassey, P., Evans, N. J., Grasman, R. P. P., et al. (2019). The quality of response time data inference: A blinded, collaborative assessment of the validity of cognitive models. *Psychonomic Bulletin & Review*, *26*(4), 1051–1069. <https://doi.org/10.3758/s13423-017-1417-2>.
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ-British Medical Journal*, *315*(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>.
- Evans, N. J. (2020). Think fast! The implications of emphasizing urgency in decision-making. *PsyArXiv*. <https://doi.org/10.31234/osf.io/pfrb4>
- Forstmann, B. U., Tittgemeyer, M., Wagenmakers, E.-J., Derrfuss, J., Imperati, D., & Brown, S. (2011). The speed-accuracy tradeoff in the elderly brain: A structural model-based approach. *Journal of Neuroscience*, *31*(47), 17242–17249. <https://doi.org/10.1523/JNEUROSCI.0309-11.2011>.
- Foster, J. C., & Taylor, G. A. (1920). The applicability of mental tests to persons over fifty years of age. *Journal of Applied Psychology*, *4*(1), 39–58. <https://doi.org/10.1037/h0071260>.
- Garton, R., Reynolds, A., Hinder, M. R., & Heathcote, A. (2019). Equally flexible and optimal response bias in older compared to younger adults. *Psychology and Aging*, *34*(6), 821–835. <https://doi.org/10.1037/pag0000339>.
- Gehanno, J. F., Rollin, L., & Darmoni, S. (2013). Is the coverage of google scholar enough to be used alone for systematic reviews. *Bmc Medical Informatics and Decision Making*. <https://doi.org/10.1186/1472-6947-13-7>.
- Habeck, C. W., & Schultz, A. K. (2015). Community-level impacts of white-tailed deer on understory plants in North American forests: a meta-analysis. *Aob Plants*. <https://doi.org/10.1093/aobpla/plv119>.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. <https://doi.org/10.2307/1164588>.
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. <https://doi.org/10.1002/sim.1186>.

- Huff, M. J., & Aschenbrenner, A. J. (2018). Item-specific processing reduces false recognition in older and younger adults: Separating encoding and retrieval using signal detection and the diffusion model. *Memory & Cognition*, *46*(8), 1287–1301. <https://doi.org/10.3758/s13421-018-0837-1>.
- Kapucu, A. (2010). *Emotional recognition memory for younger and older adults: Combining ROC analysis and the diffusion model (Dissertation)*. Amherst, MA: University of Massachusetts Amherst.
- Kordella, B. C. (2009). *A diffusion model analysis of the effects of aging on sentence memory (Master thesis)*. Athens, OH: Ohio State University.
- Kühn, S., Schmiedek, F., Schott, B., Ratcliff, R., Heinze, H.-J., Düzel, E., et al. (2011). Brain areas consistently linked to individual differences in perceptual decision-making in younger as well as older adults before and after training. *Journal of Cognitive Neuroscience*, *23*(9), 2147–2158. <https://doi.org/10.1162/jocn.2010.21564>.
- Leite, F. P., & Ratcliff, R. (2011). What cognitive processes drive response biases? A diffusion model analysis. *Judgment and Decision Making*, *6*(7), 651–687.
- Lerche, V., Christmann, U., & Voss, A. (2018). Impact of context information on metaphor elaboration: A diffusion model study. *Experimental Psychology*, *65*(6), 370–384. <https://doi.org/10.1027/1618-3169/a000422>.
- Lerche, V., von Krause, M., Voss, A., Frischkorn, G. T., Schubert, A.-L., & Hagemann, D. (2020). Diffusion modeling and intelligence: Drift rates show both domain-general and domain-specific relations with intelligence. *Journal of Experimental Psychology: General*, *1*, 1. <https://doi.org/10.1037/xge0000774>.
- Lerche, V., & Voss, A. (2016). Model complexity in diffusion modeling: Benefits of making the model more parsimonious. *Frontiers in Psychology*, *7*(1324), 1. <https://doi.org/10.3389/fpsyg.2016.01324>.
- Lerche, V., & Voss, A. (2017). Experimental validation of the diffusion model based on a slow response time paradigm. *Psychological Research Psychologische Forschung*. <https://doi.org/10.1007/s00426-017-0945-8>.
- Lerche, V., & Voss, A. (2018). Speed–accuracy manipulations and diffusion modeling: Lack of discriminant validity of the manipulation or of the parameter estimates? *Behavior Research Methods*, *50*(6), 2568–2585. <https://doi.org/10.3758/s13428-018-1034-7>.
- Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, *16*(5), 798–817. <https://doi.org/10.3758/PBR.16.5.798>.
- McKoon, G., & Ratcliff, R. (2012). Aging and IQ effects on associative recognition and priming in item recognition. *Journal of Memory and Language*, *66*(3), 416–437. <https://doi.org/10.1016/j.jml.2011.12.001>.
- McKoon, G., & Ratcliff, R. (2013). Aging and predicting inferences: A diffusion model analysis. *Journal of Memory and Language*, *68*(3), 240–254. <https://doi.org/10.1016/j.jml.2012.11.002>.
- Moreno, S. G., Sutton, A. J., Ades, A. E., Stanley, T. D., Abrams, K. R., Peters, J. L., et al. (2009). Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Medical Research Methodology*. <https://doi.org/10.1186/1471-2288-9-2>.
- R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108.
- Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review*, *9*(2), 278–291. <https://doi.org/10.3758/bf03196283>.
- Ratcliff, R. (2008). Modeling aging effects on two-choice tasks: Response signal and response time data. *Psychology and Aging*, *23*(4), 900–916. <https://doi.org/10.1037/a0013930>.
- Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, *111*(1), 159–182. <https://doi.org/10.1037/0033-295x.111.1.159>.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922. <https://doi.org/10.1162/neco.2008.12-06-420>.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347–356. <https://doi.org/10.1111/1467-9280.00067>.
- Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging*, *19*(2), 278. <https://doi.org/10.1037/0882-7974.19.2.278>.
- Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, *16*(2), 323–341.
- Ratcliff, R., Thapar, A., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on brightness discrimination. *Percept Psychophys*, *65*(4), 523–535. <https://doi.org/10.3758/bf03194580>.
- Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language*, *50*(4), 408–424. <https://doi.org/10.1016/j.jml.2003.11.002>.
- Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging and individual differences in rapid two-choice decisions. *Psychonomic Bulletin & Review*, *13*(4), 626–635. <https://doi.org/10.3758/BF03193973>.
- Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology*, *60*(3), 127–157. <https://doi.org/10.1016/j.cogpsych.2009.09.001>.
- Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and IQ on item and associative memory. *Journal of Experimental Psychology: General*, *140*(3), 464–487. <https://doi.org/10.1037/a0023810>.
- Ratcliff, R., Thompson, C. A., & McKoon, G. (2015). Modeling individual differences in response time and accuracy in numeracy. *Cognition*, *137*, 115–136. <https://doi.org/10.1016/j.cognition.2014.12.004>.
- Ratcliff, R., & Van Dongen, H. P. A. (2009). Sleep deprivation affects multiple distinct cognitive processes. *Psychonomic Bulletin & Review*, *16*(4), 742–751. <https://doi.org/10.3758/PBR.16.4.742>.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. <https://doi.org/10.1037/0033-295X.103.3.403>.
- Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, *13*(4), 140–144. <https://doi.org/10.1111/j.0963-7214.2004.00293.x>.
- Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology General*, *136*(3), 414–429. <https://doi.org/10.1037/0096-3445.136.3.414>.
- Schmitz, F., & Wilhelm, O. (2016). Modeling mental speed: Decomposing response time distributions in elementary cognitive tasks and correlations with working memory capacity and fluid intelligence. *Journal of Intelligence*, *4*(4), 13.
- Schubert, A.-L., Hagemann, D., Voss, A., Schankin, A., & Bergmann, K. (2015). Decomposing the relationship between mental speed and mental abilities. *Intelligence*, *51*, 28–46. <https://doi.org/10.1016/j.intell.2015.05.002>.
- Shultz, M. (2007). Comparing test searches in PubMed and google scholar. *Journal of the Medical Library Association*, *95*(4), 442–445. <https://doi.org/10.3163/1536-5050.95.4.442>.

- Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(1), 101–117. <https://doi.org/10.1037/0278-7393.32.1.101>.
- Spaniol, J., Voss, A., Bowen, H. J., & Grady, C. L. (2011). Motivational incentives modulate age differences in visual perception. *Psychology and Aging*, 26(4), 932–939. <https://doi.org/10.1037/a0023297>.
- Spaniol, J., Voss, A., & Grady, C. L. (2008). Aging and emotional memory: Cognitive mechanisms underlying the positivity effect. *Psychology and Aging*, 23(4), 859–872. <https://doi.org/10.1037/a0014218>.
- Sterne, J. A., & Egger, M. (2005). Regression Methods to Detect Publication and Other Bias in Meta-Analysis. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication Bias in Meta-Analysis* (pp. 99–110). Chichester, UK: John Wiley & Sons Ltd.
- Thapar, A., Ratcliff, R., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychology and Aging*, 18(3), 415–429. <https://doi.org/10.1037/0882-7974.18.3.415>.
- van Ravenzwaaij, D., Donkin, C., & Vandekerckhove, J. (2017). The EZ diffusion model provides a powerful test of simple empirical effects. *Psychonomic Bulletin & Review*, 24(2), 547–556. <https://doi.org/10.3758/s13423-016-1081-y>.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor Package. *Journal of Statistical Software*, 36(3), 1–48.
- Viechtbauer, W. (2018, December 8). I^2 for Multilevel and Multivariate Models. Retrieved from: https://www.metafor-project.org/doku.php/tips:i2_multilevel_multivariate/
- Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2), 112–125. <https://doi.org/10.1002/jrsm.11>.
- Voskuilen, C., Ratcliff, R., & McKoon, G. (2018). Aging and confidence judgments in item recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(1), 1–23. <https://doi.org/10.1037/xlm0000425>.
- Voss, A., Lerche, V., Mertens, U., & Voss, J. (2019). Sequential sampling models with variable boundaries and non-normal noise: A comparison of six models. *Psychonomic Bulletin & Review*, 26(3), 813–832. <https://doi.org/10.3758/s13423-018-1560-4>.
- Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology: a practical introduction. *Exp Psychol*, 60(6), 385–402. <https://doi.org/10.1027/1618-3169/a000218>.
- Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: an empirical validation. *Mem Cognit*, 32(7), 1206–1220.
- Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, 21(5), 641–671. <https://doi.org/10.1080/09541440802205067>.
- Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, 58(1), 140–159. <https://doi.org/10.1016/j.jml.2007.04.006>.
- Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, 14(1), 3–22. <https://doi.org/10.3758/bf03194023>.
- Willoughby, R. R. (1929). Incidental learning. *Journal of Educational Psychology*, 20(9), 671–682. <https://doi.org/10.1037/h0071404>.
- Yap, M. J., Sibley, D. E., Balota, D. A., Ratcliff, R., & Rueckl, J. (2015). Responding to nonwords in the lexical decision task: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(3), 597–613. <https://doi.org/10.1037/xlm0000064>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.