2025 ■ Vol. 21 ■ no. 3

A replication of Ratcliff and Hacker's (1981) Same-Different task variant (Experiment 1)

Silia Losier Poirier [a], Audrey Jessica Caissie [a] & Bradley Harding [a]

[a]Université de Moncton

Abstract ■ The *Same-Different* task consists of a rapid presentation of two stimuli and is used to assess how quickly humans can distinguish between identical and non-identical stimuli. Despite the task's simplicity and benefitting from over 50 years of research, there are still major gaps in the models that try to explain the cognitive process required to make comparison judgements. A central modelling hurdle is the *fast-same* phenomenon, referring to the fact that "Same" responses are faster and more accurate than "Different" responses, whereas typical analytical-based models would predict the opposite. Nevertheless, some modelling propositions have been put forth, including Ratcliff and Hacker's (1981) response bias framework. In this study, we diligently reproduced Experiment 1 of Ratcliff and Hacker's (1981) study to put their proposition to the test. This experimental study consists of two conditions: adding a caution component to "Same" responses in one case and to "Different" responses in the other. We observe that manipulating task instructions does not generate any significant difference between conditions and that there is no speed-accuracy trade-off, opposing Ratcliff and Hacker's (1981) original results. Response bias could therefore not be the main explanation for the fast-same phenomenon.

Keywords ■ Same-Different task, fast-same effect.

✉ esl4535@umoncton.ca

10.20982/tqmp.21.3.p115

Acting Editor ■ Denis Cousineau (Université d'Ottawa)

Reviewers
■ No authors from the original article reviewed this article na

Introduction

The *Same-Different* task is a decision-making task in which individuals are presented with two stimuli, and it is used to assess how similarity/difference(s) is/are detected between items. In the classic implementation of the task, participants are required to judge as quickly and as accurately as possible whether these successively presented stimuli were entirely identical or not. Even if the task is simple for participants to complete, the decision-making processes involved remain unclear. Consensus has not yet been reached among researchers, as existing models are unable to predict the *fast-same* phenomenon: individuals tend to respond faster and more accurately to identical pairs of stimuli (Bamber, 1972) than what models can predict (Harding & Cousineau, 2022).

With the task finding its roots in the cognitive boom of the 1950s and 1960s, several researchers have since suggested models that could explain the task (e.g., Bamber, 1969; Krueger, 1978; Nickerson, 1976; Taylor, 1976), including Robert Proctor's encoding facilitation approach (Proctor, 1981) and Roger Ratcliff's response bias approach (Ratcliff & Hacker, 1981). From these opposing modelling propositions, a spirited debate ensued (in chronological order: Proctor, 1981; Ratcliff & Hacker, 1981; Proctor & Rao, 1982, 1983, 1983, 1984, 1985, 1986 where no concrete conclusion emerged and no satisfying proposition was made; modelling the task came to a stalemate for 35 years.

Nevertheless, there has been recent progress in understanding the cognitive processes involved when comparing stimuli, progress which was made by revisiting past hypotheses on identity priming and response bias (Cousineau et al., 2023; Goulet & Cousineau, 2020; Harding & Cousineau, 2022). From this renewed interest, Harding and Cousineau (2022) were able to validate identity priming as a possible explanation for the fast-same phenomenon, heavily leveraging Proctor's (1981) methods and results in their four Same-Different task variants. However, while they identified identity priming as the likely process responsible for the fast-same phenomenon (a find-

ing supported by Cousineau et al., 2023), as they point out, a small advantage for "Same" responses remains when priming is cancelled (or attenuated). Therefore, a replication of Ratcliff and Hacker (1981) study, which rejects this explanation, is warranted to examine their response bias hypothesis, which has not received as much attention.

Ratcliff and Hacker (1981) Experiment 1

The main objective of Ratcliff and Hacker's (1981) Experiment 1 was to examine the effect of criterion change (taking advantage of the speed-accuracy tradeoff, first introduced by Henmon, 1911; see Heitz, 2014, for a review) in response time (RT) within a comparison task construct. Their study consisted of two experiments in which they used a series of four-letter strings as stimuli.

In Experiment 1, they manipulated the instructions given to participants to add a *caution* component. In the first condition, participants must only answer "Same" if they are absolutely certain of their answer. In the second condition, they must answer "Different" only if they are absolutely certain of their answer. They had four participants complete three sessions for both cautious conditions (4 blocks $\times$ 96 trials $\times$ 3 sessions = 1152 data/participant/-condition with 576 trials/condition for "Same" and 576 trials/condition for "Different"). Their results, seen in Table 1, showed that in the cautious-"Same" condition, RT were slower for "Same" stimulus pairs. Furthermore, participants' responses were more accurate for "Same" stimuli in this condition. In the cautious-"Different" condition, "Different" pairs were less accurate, except for pairs of stimuli differing by four letters.

Ratcliff and Hacker (1981) therefore concluded that the differences in RT between "Different" and "Same" are manipulable – that the fast-same phenomenon is simply the result of a preference for *sameness*. With this conclusion, they thus rejected priming and identity reporting models (Bamber, 1969; Proctor, 1981; Proctor & Rao, 1982, 1983; Proctor et al., 1984; Proctor, 1986; Taylor, 1976). They did, however, agree with sequential feature sampling models, such as those suggested by Krueger (1978) and Ratcliff (1981), which predict that an individual compares the two stimuli on the number of common or different characteristics.

This study

The objective of the following study is to test the response bias model by replicating Ratcliff and Hacker's (1981) Experiment 1. We believe that even if we see a change in RT between decisions for both experimental conditions, the results will be mitigated at best, considering that Ratcliff and Hacker's (1981) focused solely on four-letter stimuli, which typically have similar mean RTs between "Same" and "Different" responses.[1] We believe that we will show that response bias, manipulated through the task's instructions, is not sufficient to explain the fast-same phenomenon alone.

Methodology

This study directly replicated the methodology of Ratcliff and Hacker's (1981) Experiment 1 with the available information in the article, with a larger group of participants (rather than having 4 participants repeat the study 3 times). All details, unless otherwise specified, are as faithfully adapted as possible.

Participants and Recruitment

A total of 40 participants took part in this study, all of them over the age of 18. The first and second conditions had 26 and 25 participants, respectively, with 11 subjects having done both. A priori power analyses with a $\mathrm{B} = 0.8$, $\alpha = 0.05$, and an effect size of .597 & .093 revealed a necessity of 1-3 participants considering repeated measures[2] (Goulet & Cousineau, 2019).[3] We therefore decided to gather many more participants ($n = 20$) to match modern conventions of Same-Different task research (see Table 1 of Cousineau et al., 2023) and reduce error bar size to make appropriate and informed visual comparisons/inference of the observed effects.

The only exclusion criterion was the inability to recognise the 26 letters of the Roman alphabet and normal or

[1] A four-letter stimulus that is entirely "Same" and an entirely "Different" from its criterion stimulus have RTs that are similar with error bars that markedly overlap. See Harding and Cousineau (2022) as well as Cousineau et al. (2023), for many variants of the Same-Different task showing these error bars that are constructed with up to 5 times more participants and 4 times more total data.

[2] In this case, we calculated the expected effect size for the mean RT of "Same" responses between the cautious-"Same" condition vs. cautious-"Different" conditions using the descriptive statistics provided on p. 305 of Ratcliff and Hacker's (1981) article: mean 1 = 573 ms with an SE of 4 ms, mean 2 = 472 ms with an SE of 3 ms. We are interested in this effect size as the crux of the issue relies on whether "Same" responses are sensitive to changes in task instructions. When doing an identical analysis for "Different" responses across conditions, we observed a much smaller estimated effect size of 0.093 (conservatively estimating standard errors to be 15 ms based on metrics reported on p. 305 of Ratcliff and Hacker (1981).

[3] Considering repeated measures, we know that each participant in the original article performed both response conditions 576 times; however, no information is provided on the average correlation across measures. Nevertheless, past Same-Different task research (Goulet & Cousineau, 2019), including Appendix B of Goulet and Cousineau's (2019) article, reports a typical repeated-measure correlation of 0.2 for this task. We therefore selected a more conservative correlation of $r = 0.15$ despite reported standard errors being from 3-18 ms (Ratcliff & Hacker, 1981) for 4 participants doing the task 3 times apiece (likely increasing repeated measure correlation; a fast participant is likely always fast, and vice versa). This calculation resulted in a rounded-up required sample size of one single participant to observe the expected effect between "Same" conditions, and 3 participants to observe the expected effect between "Different" conditions.

Table 1 ■ Ratcliff and Hacker's (1981) results for response times and accuracy rates by stimulus pairs and condition.

Condition	nDiff	N	Mean RT (ms)	Mean accuracy
	0	2,303	573	.89
	1	576	605	.80
cautious-"Same"	2	576	517	.96
	3	576	480	.98
	4	576	461	.97
	0	2,303	472	.97
	1	574	695	.65
cautious-"Different"	2	576	584	.89
	3	575	531	.94
	4	576	518	.97

Note. Adapted from "Speed and accuracy of same and different responses in perceptual matching", by R. Ratcliff and M. J. Hacker, 1981, *Perception & Psychophysics*, 30(3), p. 305

corrected-to-normal vision. All participants gave written and oral consent, with the possibility of withdrawing at any moment. There was no monetary incentive. Participants who were enrolled at Université de Moncton with a class that allowed it were given 1 SONA point as compensation (a bonus credit that can be applied to the final mark of any of their participating classes).

Stimuli

The participants were presented with a criterion stimulus (S1), followed by a test stimulus (S2), on a computer screen, and they had to determine whether the two were identical or different. The stimuli were presented on a CRT computer screen with a refresh rate of 85 Hz.

Each stimulus consisted of a series of 4 letters randomly selected from: C, D, F, H, J, K, L, R, S, and T, generated and presented by E-Prime v.3.0.0.80. These letters, used in Ratcliff and Hacker's (1981) study, were used by convention as they were identified as being minimally confusable (Taylor, 1976). The letters only appeared once per stimulus. If the second stimulus was different than the first, one to four of the letters are replaced by remaining letters from the list. For the letters that remained identical, they were kept in identical positions between stimuli.

Procedure

Figure 1 shows the experimental design. The stimulus criterion (S1) appeared on the screen for one second (1000 ms). Next, an empty screen was shown for 200 ms, followed by a mask ($$$$) for 100 ms, followed by the test stimulus (S2) for 200 ms. After this, another blank screen appeared until the participant had made their decision on whether the stimuli were "Same" or "Different" by pressing the `CTRL` and `ENTER` keys on either side of the keyboard. For half the subjects, the keys for each decision were reversed to minimise possible biases that could be induced by the dominant hand pressing "Same" or "Different". After giving their answer, feedback appeared on the screen for 500 ms, showing either a plus (+) or minus (-) sign if the answer was correct or incorrect, respectively. Then another blank screen was shown for 1500 ms, and the next trial began.

The experiment consisted of two conditions, differing only in the instructions given prior to the testing phase. In the first condition, participants were instructed to only press the key for "Same" when they were *absolutely certain of their answer*. In the second condition, they were instructed to only answer "Different" when they were *absolutely certain of their answer*.

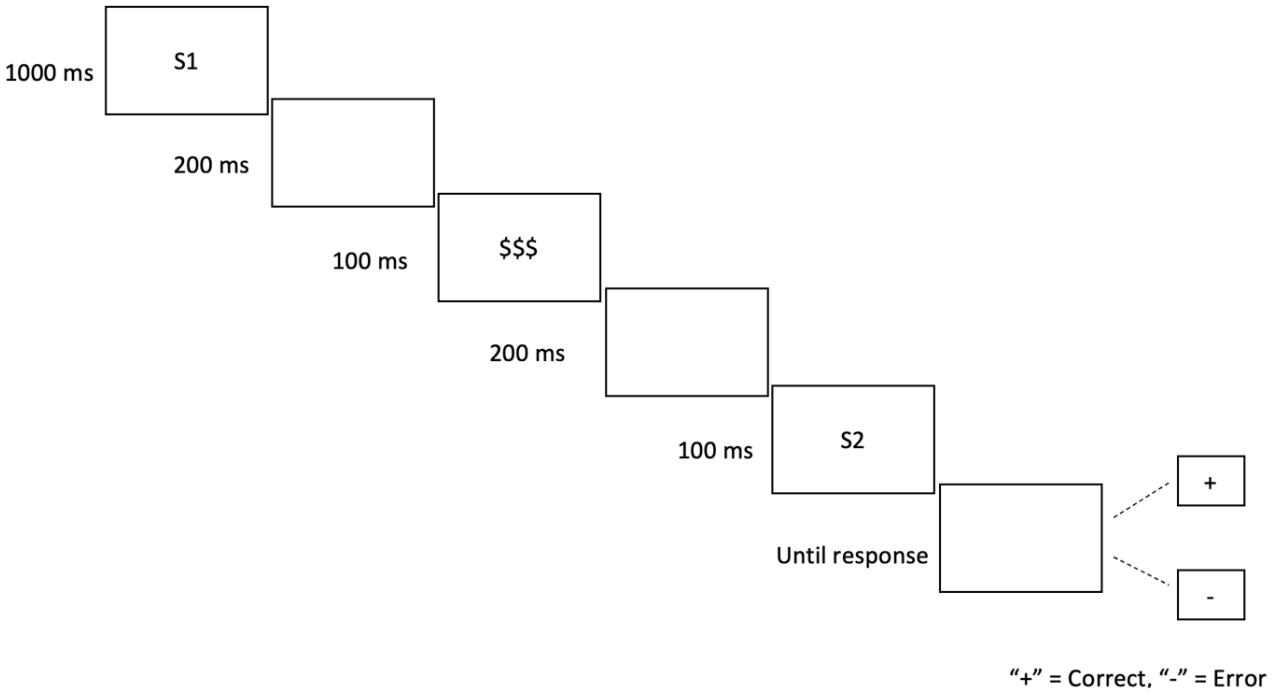
Experimental design

Each condition contained 384 experimental trials (4 blocks $\times$ 96 trials). Participants began with 16 practice trials after receiving instructions and seeing an example of a trial. They were given the opportunity to take three breaks throughout the completion of the task to minimise possible fatigue. Half of the trials consisted of identical stimuli, and the other half of different stimuli (192 "Same" trials/-condition; 192 "Different" trials/condition with 1, 2, 3 or 4 differences, having 48 trials apiece per condition).

Results

With data stemming from 26 participants in the first condition and 25 participants in the second condition, we obtained 9,984 and 9,600 RTs, respectively. The data from one participant who took part in both conditions was excluded from the analysis because their accuracy rates were abnormally low (total average of 9% accuracy), later explained by an error in the experiment's code that provided erroneous feedback to the participant.

Figure 1 ■ Schematic representation of the experimental design



We filtered the data, keeping responses with RTs above 200 milliseconds and below 2500 milliseconds (matching modern Same-Different task conventions put forth in Harding and Cousineau (2022), which excluded 6 and 160 RTs, respectively. We therefore have a total of 18,650 RTs to analyse; 9,521 for Condition caution-"Same" and 9,129 for Condition caution-"Different", resulting in 0.017% rejection. We also analysed the data from 10 participants who had completed both conditions of the task, which resulted in 7,605 total data points for within-subject analyses.

Accuracy

Table 2 presents the overall mean RT and accuracy for "Same" and "Different" responses for each condition. It also presents the RTs and accuracy based on the number of letters differing between stimulus pairs. There were 530 errors in Condition caution-"Same" and 626 errors in Condition cautious-"Different". Therefore, there was a global error rate of 5.57% and 6.86%, respectively. For both conditions, average accuracy is very similar, although slightly higher for "Same" pairs. When examining the average for each number of differences, accuracy is lower for "Same" responses than for most "Different" responses, except for the lowest accuracy being for pairs differing by a single letter; this is true for both conditions and typical of Same-Different task results (see Harding & Cousineau, 2022).

Response times

For RT analyses, only correct responses were analysed (Condition cautious-"Same", $n = 8,991$; Condition cautious-"Different", $n = 8,503$). Table 2 shows that for both conditions, the average RT for "Same" pairs is slightly slower than for "Different" pairs. Furthermore, Table 2 shows that, overall, pairs of stimuli differing by three or four letters had the highest accuracy rates and fastest RT; conversely, single-difference stimuli produced the lowest performance for both metrics.

Our mean RT, mean accuracy, mean standard deviations, and mean asymmetry, are shown by the number of differing letters between stimulus pairs in Figure 2. The error bars represent the difference- and correlation-adjusted 95% confidence intervals (CI; Cousineau et al., 2021); each CI is constructed using the appropriate standard error and CI estimator (Harding et al., 2014). The distributions are clearly very similar for both conditions. As can be seen, stimuli with a longer average RT also have lower accuracy, which is typically observed (Cousineau et al., 2023; Harding & Cousineau, 2022)=.

Table 2 ■ Descriptive statistics for response times and accuracy rates, by stimulus pairs "Same" and "Different", and then by number of differing letters, according to condition.

Condition	Stimulus	RT (ms)			Accuracy rates		
		N	Mean	SD	N	Mean	SD
	Same	4,541	660	318	4 760	.95	.21
	Different	4,450	643	284	4 761	.93	.25
cautious-	1D	975	748	338	1 179	.83	.38
"Same"	2D	1,132	661	293	1 190	.95	.21
	3D	1,173	596	236	1 197	.98	.14
	4D	1,170	584	240	1 195	.98	.14
	Same	4,291	635	285	4 568	.94	.24
	Different	4,212	629	282	4561	.92	.27
cautions-	1D	924	729	316	1 138	.81	.39
"Different"	2D	1,079	644	275	1 140	.95	.22
	3D	1,109	596	273	1 144	.97	.17
	4D	1,100	564	238	1 139	.97	.18

Note. D = Number of differing letters in a pair of stimuli.

Between-Subject Analysis

To contrast between the results of Condition cuatious-"Same" and Condition cautious-"Different", we performed a factorial ANOVA with one factor for the response modality ("Same"/"Different" as a within-subject factor) and a between-subject factor for the experimental condition (Condition cautious-"Same"/Condition cautious-"Different"). To avoid any contamination that could be related to subjects who participated in both conditions, this analysis is carried out solely on those who only participated in Condition cautious-"Same" ($n = 15$), and those who only participated in Condition cautious-"Different" ($n = 14$); this subsample remains well within the statistical power requirements. We discover that there is no significant interaction between factors, $F(1, 27) = 1.023$, $p = .321$, $\eta_p^2 = .036$, no significant main effect on the between-subject factor (experimental condition), $F(1, 27) = .054$, $p = .818$, $\eta_p^2 = .002$, and no significant main effect on the within-subject factor (response modality), $F(1, 27) = .540$, $p = .469$, $\eta_p^2 = .020$.

Within-Subject Analysis

In the original study, Ratcliff and Hacker (1981) had the participants perform both conditions of the tasks, which could lead to noteworthy results here, especially considering the power that replicated measures have on the overall results (Goulet & Cousineau, 2019). Therefore, we performed a 2× 2 within-subject ANOVA solely on subjects who participated in both conditions. This ANOVA has one factor for the response modality ("Same"/"Different") and another factor for the experimental condition (Condition cautious-"Same"/Condition cautious-"Different"). We observe no significant interaction between factors, $F(1, 9) = .959$, $p = .353$, $\eta_p^2 = .096$, no significant main effect for the experimental condition factor, $F(1, 9) = .958$, $p = .353$, $\eta_p^2 = .096$, and no significant main effect for the response modality factor, $F(1, 9) = .399$, $p = .543$, $\eta_p^2 = .043$.

Table 3 shows the correlations between the RTs of Condition cautious-"Same" and Condition cautious-"Different" for these 10 participants, based on the number of differing letters between the two stimuli. There is a strong correlation from one condition to another; participants who were fast in one condition were fast in other conditions as well.

EZ diffusion model

We modelled the results using the EZ diffusion model (Wagenmakers et al., 2007), and the model predictions are consistent with those presented in Harding and Cousineau (2022), Appendix B. Table 4 presents the parameter estimates for accumulation rate, threshold, and non-response time (v, a, and T_{er} respectively), which are very similar for both conditions.

As was the case in Harding and Cousineau (2022) article on the Same-Different task, we observe mainly a change in the non-response time parameter rather than the threshold parameter. If response bias had been the major factor in the changes in RT within this task, the latter would have been expected to vary significantly in relation to the response modality.

Discussion

This study aimed to directly replicate Ratcliff and Hacker's (1981) Experiment 1 as faithfully as possible to test the re-

Table 3 ■ Correlation matrix of response times for participants who completed both conditions.

	Condition caution-"Different"				
Condition caution-"Same"	Same	1D	2D	3D	4D
Same	.838*	.849*	.732**	.731**	.803*
1D		.718**	.780*	.805*	.872*
2D			.716**	.811*	.815*
3D				.775*	.809*
4D					.897*

Note. $N = 10$. *: $p < .01$; **: $p < .05$

Table 4 ■ EZ model estimation of parameters v, T_{er}, and a according to stimulus pairs and condition.

Condition	$nDiff$	v	T_{er}	a
	0	0.000230	332	0.152
	1	0.000140	406	0.134
cautious-"Same"	2	0.000260	343	0.162
	3	0.000333	325	0.165
	4	0.000331	325	0.157
	0	0.000219	323	0.141
	1	0.000128	375	0.134
cautious-"Different"	2	0.000244	332	0.153
	3	0.000282	268	0.175
	4	0.000305	298	0.145

Note. nDiff = the number of different letters in between the pair of stimuli; v = accumulation rate; T_{er} = residual time; a = threshold.

sponse bias framework, as it is one of the only published articles to propose an alternative to the identity priming model put forth by Proctor (1981; concretised in 2022). Our hypotheses, based on the results and conclusions of previous empirical studies (Cousineau et al., 2023; Proctor, 1986; Proctor & Rao, 1982, 1983), were that 1) that we would obtain a mitigated at -best effect of response bias when manipulating task instructions to take advantage of the speed-accuracy trade-off, and 2) that there would be no significant trade-off between RT and accuracy. The results obtained support both hypotheses. Mean RTs from both conditions are similar, indicating that a caution-instruction component influencing response bias has a negligible effect on RTs and accuracy.

These results contrast with those of Ratcliff and Hacker (1981), and the present study's mean RT patterns (shown in Figure 2 and described in Table 2) converge with modern variants of the Same-Different task (Harding & Cousineau, 2022). Nevertheless, there are minor changes between conditions that require assessment. First, for the participants who took part in both conditions, the cautious-"Same" condition has slower mean RT for "Same" responses as their cautious-"Different" counterparts (a difference of 12 ms; n.s.), second, there was a difference between the mean RTs of both "Different" conditions (a difference of 32 ms; n.s.), third, there was no significant difference between the overall mean RTs of "Same" and "Different" in Condition cautious-"Same" (a difference of 17 ms; n.s.), fourth, there were no significant differences between the overall mean RTs of "Same" and "Different" responses in Condition cautious-"Different" (a difference of 6 ms, n.s.). We consistently observed small to negligible effect sizes for both between- and within-subject analyses, whereas the expected metric for the fast-same phenomenon (elaborated and developed in Harding & Cousineau, 2022) is a Hedges' G of = .5, and of = .15 in attenuated-"Same" conditions (when priming is removed). This may be in part due to masking the criterion stimulus (as is done in this replication and in the original article), which has been shown to reduce/cancel residual activation within the neuronal pathways (Huber, 2008). This hindered residual activation is at the centre of the attenuated-"Same" of Harding and Cousineau's (2022) proposition, and it is therefore unsurprising to see such a small discrepancy between "Same" and "Different" mean RTs here; there was no baseline fast-same phenomenon within this experiment, and that is likely due to masking. Additionally, the present experiment cannot conclude that a fast-same phenomenon was ob-

served, especially considering that four-letter stimuli typically have error bars for all-“Same” and all-“Different” that typically entirely overlap (Harding & Cousineau, 2022). Furthermore, while there are discrepancies between the overall mean RTs of “Same” and “Different” responses, this is expected as “Same” responses typically have lower intercepts than those for “Different” Responses (Harding & Cousineau, 2022). The results that we find are therefore unsurprising.

In the original experiment, all participants took part in both conditions. While we focused on gathering a larger sample size with fewer replicated measures per participant (384 trials/participant here, 1152 trials/participant in the original article) , this change should have minimal effect considering power requirements noted above and that our distributions are constructed with more between-participant variability – a methodological change that would show more of the effect across a sample rather than have error bars be decreased by decree of having less within-subject variability (as noted, participants who are typically fast in one condition, are also fast in all conditions, see Table 3). Nevertheless, with our 10 subjects who participated in both conditions, we found no significant differences at any level.

In the original article, the authors concluded that there was a speed-accuracy trade-off, yet they obtained the highest accuracy rates for stimuli with the fastest response times; we also obtained this result. The original article also put emphasis on mean RT and accuracy alone, without analysing variance or correlation to assess their participants’ performance stability. There were also no inference tests to assess potential significant differences between conditions. Our study therefore allows for a more rigorous evaluation of the effect of manipulating the decision criterion.

However, the question persists: why was Ratcliff and Hacker’s (1981) experiment important to replicate when recent advancements have carried momentum in an orthogonal direction? We maintain that it is because their noteworthy results are convincing of their case; the effect size that they report is so high, it demands a second look to observe how the only serious contender to the priming hypothesis may operate. While it is evident that we are proponents of the identity priming approach to modelling the Same-Different task (the supervising author has published, or participated in the few recently published articles addressing this niche moment in cognitive science history), replication of a competing proposition allows us to move forward all while being informed ? it ensures that we are not leaving data on the table that could give a better understanding of the processes at play.

Nevertheless, now that we have these results, we argue that we do not replicate the original article’s results for two interrelated reasons: 1) their choice of a within-subject design to have more data by fewer participants (a large m, small n design which typically produces robust results, Smith & Little, 2018), and 2) the selection of their participants. Considering the first reason, we have discussed how the Same-Different task can benefit from many data per participant because the intra-participant correlation is expected to be high, i.e., participants who are fast in one condition, are fast in all conditions (see Table 3 for our empirical results attesting to this; the intraclass participant correlation for the present article is $r = 0.33$). In this case, between-subject variability is lost and mean data become more reliant (and representative) of single participant performances. Considering the second reason, Ratcliff and Hacker (1981) describe participants on p. 305. They are the paper’s second author (Michael J. Hacker), a Rockefeller University faculty volunteer, and two compensated local long-term subjects. It could therefore be argued that this selection makes for participants who are likely to perform well on cognitive tasks as they are either a) involved in cognitive research, or b) used to participating in cognitive research. By combining the two enumerated reasons, we can infer why the original paper’s effect was as high as it was: a few motivated participants performed very well at identifying “Same” and “Different” stimuli because they are used to these types of tasks, and they performed this task many times. While the experimental design choices of the original article typically result in robust findings, we believe this may be a case for not seeing the forest for the trees ? that the salient effect size led to a response-bias conclusion that was not generalisable and only representative of the original paper’s results.

Conclusion

Ratcliff and Hacker’s (1981) work is important to the storied modelling debate surrounding the Same-Different task, and the response bias framework remained important to replicate as it offered what all science needs to move forward: a model requiring falsification. We do not contradict that decision-making processes are innately complex to decipher and that response biases may be present in any of our perceptive/cognitive processes (in fact, the intercept effect noted in Harding & Cousineau, 2022, could be interpreted as a response bias). We do, however, conclude that Ratcliff and Hacker’s (1981) response bias proposition is unlikely to be the sole cause of the fast-same phenomenon, but rather a complement (at best) to the response facilitation approach (Harding & Cousineau, 2022; Proctor, 1981).

This conclusion will be concretised by carrying out a replication of Ratcliff and Hacker (1981) Experiment 2, in which letters between stimuli are transposed from one po-

sition in the stimulus string to another; an experiment originally used to supplement Experiment 1's now irreplicable findings.

Authors' note

This research has been funded by*Research NB* (formerly the *New Brunswick Innovation Fund*), an *Undergraduate Student Research Award* from the *National Science Research and Engineering Council of Canada*, attributed to the second author. This research was also supported by internal funding from the Université de Moncton's research office (FESR). We would like to thank Thomas Pelletier and Josée Sparks for their helpful comments on a previous version of the manuscript..

References

Bamber, D. (1969). Reaction times and error rates for "same" and "different" judgments of multidimensional stimuli. *Perception & Psychophysics*, *6*(3), 169–174. doi: 10.3758/BF03210087.

Bamber, D. (1972). Reaction times and error rates for judging nominal identity of letter strings. *Perception & Psychophysics*, *12*(4), 321–326. doi: 10.3758/BF03207214.

Cousineau, D., Goulet, M.-A., & Harding, B. (2021). Summary plots with adjusted error bars [Article 25152459211035109]. *The superb framework with an implementation in R. Advances in Methods and Practices in Psychological Science*, *4*(3). doi: 10 . 1177 / 25152459211035109.

Cousineau, D., Harding, B., Walker, J. A., Durand, G., T.-Groulx, J., Lauzon, S., & Goulet, M.-A. (2023). Analyses of response time data in the same–different task. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale*, *77*(2), 115–129. doi: 10.1037/cep0000301.

Goulet, M.-A., & Cousineau, D. (2019). The power of replicated measures to increase statistical power. *Advances in Methods and Practices in Psychological Science*, *2*(3), 199–213. doi: 10.1177/2515245919849434.

Goulet, M.-A., & Cousineau, D. (2020). Sequential sampling models of same-different data and how they explain the fast-same effect. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale*, *74*(4), 284–301. doi: 10.1037/cep0000197.

Harding, B., & Cousineau, D. (2022). Is the fast-same phenomenon that fast? an investigation of identity priming in the same–different task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(4), 520–546. doi: 10.1037/xlm0001076.

Harding, B., Tremblay, C., & Cousineau, D. (2014). Standard errors: A review and evaluation of standard error estimators using monte carlo simulations. *The Quantitative Methods for Psychology*, *10*(2), 107–123. doi: 10.20982/tqmp.10.2.p107.

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience*, *8*(150), 1–19. doi: 10.3389/fnins.2014.00150.

Henmon, V. A. C. (1911). The relation of the time of a judgment to its accuracy. *Psychological Review Review*, *18*, 186–201. doi: 10.1037/h0074579.

Huber, D. E. (2008). Immediate priming and cognitive aftereffects. *Journal of Experimental Psychology: General*, *137*(2), 324–347. doi: 10.1037/0096-3445.137.2.324.

Krueger, L. E. (1978). A theory of perceptual matching. *Psychological Review*, *85*(4), 278–304. doi: 10.1037/0033-295X.85.4.278.

Nickerson, R. S. (1976). Short-term retention of visually presented stimuli: Some evidence of visual encoding. *Acta Psychologica*, *40*, 153–162. doi: 10.1016/0001-6918(76)90010-5.

Proctor, R. W. (1981). A unified theory for matching-task phenomena. *Psychological Review*, *88*(4), 291–326. doi: 10.1037/0033-295X.88.4.291.

Proctor, R. W. (1986). Response bias, criteria settings, and the fast-same phenomenon: A reply to ratcliff. *Psychological Review*, *93*(4), 473–477. doi: 10.1037/0033-295X.93.4.473.

Proctor, R. W., & Rao, K. V. (1982). On the "misguided" use of reaction-time differences: A discussion of ratcliff and hacker (1981). *Perception & Psychophysics*, *31*(6), 601–604. doi: 10.3758/BF03204200.

Proctor, R. W., & Rao, K. V. (1983). Evidence that the same-different disparity in letter matching is not attributable to response bias. *Perception & Psychophysics*, *34*(1), 72–76. doi: 10.3758/BF03205898.

Proctor, R. W., Rao, K. V., & Hurst, P. W. (1984). An examination of response bias in multiletter matching. *Perception & Psychophysics*, *35*(5), 464–476. doi: 10.3758/BF03203923.

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, *88*(6), 552–572. doi: 10.1037/0033-295X.88.6.552.

Ratcliff, R. (1985). Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychological Review*, *92*(2), 212–225. doi: 10.1037/0033-295X.92.2.212.

Ratcliff, R., & Hacker, M. J. (1981). Speed and accuracy of same and different responses in perceptual matching. *Perception & Psychophysics*, *30*(3), 303–307. doi: 10.3758/BF03214286.

Ratcliff, R., & Hacker, M. J. (1983). On the misguided use of reaction-time differences: A reply to proctor & rao

(1982). *Perception & Psychophysics*, *31*(6), 603–604. doi: 10.3758/BF03204201.

Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-n design. *Psychonomic Bulletin and Review*, *25*, 2083–2101. doi: 10.3758/s13423-018-1451-8.

Taylor, D. A. (1976). Effect of identity in the multiletter matching task. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(3), 417–428. doi: 10.1037/0096-1523.2.3.417.

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. doi: 10.3758/BF03194023.

Citation

Losier Poirier, S., Caissie, A. J., & Harding, B. (2025). A replication of Ratcliff and Hacker's (1981) Same-Different task variant (Experiment 1). *The Quantitative Methods for Psychology*, *21*(3), 115–124. doi: 10.20982/tqmp.21.3.p115.

Received: 07/11/2025 ~ Accepted: 12/12/2025

Figure 2 ■ Representation of the distribution for response times, accuracy, standard deviations, and asymmetry according to the number of letters of difference for Conditions 1 and 2. The left column represents data for Condition cautious-"Same", and the right column represents data for Condition cautious-"Different". The blue bars are for "Same" stimuli. The yellow, purple, red, and black bars are for stimuli that differ by 1, 2, 3, and 4 letters, respectively. The error bars represent the difference and correlation-adjusted 95% confidence intervals (Cousineau et al., 2021)