Investigation of the Cognitive Mechanisms of Same and Different Judgments

Marc-André Goulet

Thesis submitted to the University of Ottawa
In partial fulfillment of the requirements for the
Ph.D. in experimental psychology

School of Psychology
Faculty of Social Sciences
University of Ottawa

Acknowledgments / Remerciements

The research projects included in this thesis were partially possible thanks to the financial support of the *Fonds de recherche du Québec – Nature et technologies,* of the *Ministry of College and Universities of Ontario*, of the *School of Psychology of the University of Ottawa* and of the *Pierre-Baron memorial scholarship*.

À tous ceux et celles qui m'ont supporté et aidé durant la rédaction de cette thèse, un grand merci du fond du cœur. La recherche scientifique est difficile, éprouvante et exigeante. J'ai eu la grande chance de pouvoir compter sur une famille qui m'a soutenu durant ce long processus. À ma conjointe, Camille, merci d'avoir été mon oreille et mon épaule. À ma mère, Rose-Marie, mon père, Marco, et ma sœur cadette, Alexandra, merci pour votre support depuis le début de mes études universitaires.

Je tiens également à remercier tous les collègues que j'ai côtoyés durant les dernières années. Je suis nostalgique des profondes discussions que nous avons eues. Merci à Bradley Harding pour ses conseils et son mentorat. Merci aux membres du laboratoire Vision-Integration-Cognition et du feu Quibb pour leur rétroaction toujours pertinente sur mes projets. Finalement, merci à mon superviseur Denis Cousineau pour son encadrement, pour la confiance qu'il m'a accordée et pour avoir fait de moi le chercheur que je suis devenu.

Abstract

The Same-Different task is an experimental paradigm in which a stimulus pair is presented in succession to a participant whose task is to determine if the stimuli are *Same* or *Different*. Typical results show that participants tend to be quicker to respond *Same* then they are to respond *Different*. Since the 1960s, many models were proposed to explain this effect, but none has yielded conclusive evidence. The objective of this thesis is to test these models with three experiments by focusing on three research questions: 1) what is the source of the effect, the participant or the stimuli?; 2) what is the organization of the cognitive mechanisms underlying the task?; and 3) what is the effect of the number of attributes on the processing capacity? Results show that the fast-same effect stems from the characteristics of the stimuli rather than an inherent preference for sameness. They also show that the cognitive architecture underlying the task is serial, but that it does not seem to explain solely the fast-same effect. Indeed, the fast-same effect seems to be rather caused by a more efficient processing of *Same* stimuli in the first 500 ms of the treatment compared to *Different* stimuli.

<u>Keywords:</u> Same-Different task, Fast-same effect, Cognitive architecture, Workload capacity

Table of content

Prologue

Time is critical to understand fundamental cognitive behaviours. A cognitive psychologist is a timekeeper with a plan. At its core, this thesis reports how fast humans pressed a button on a computer keyboard 137,440 times. This is about one eighth of the number of trials that Jeremy Wolfe used to conclude that "*your favorite theory of visual search is wrong [… and] so is mine.*" (Wolfe, 1998, p. 38). It puts my efforts to answer a simple question in perspective.

The research question that thrives this thesis, and one that I have come to ask myself constantly over the past eight years, is why are humans faster to respond *Same* than they are to respond *Different*? The earliest I could find someone asking themselves this question was 1964, and this person is Raymond S. Nickerson.[1] He might also be the first, since he does not cite any other publications in his paper. From what I understand from this article, Nickerson was conducting a large study on response times in various cognitive experiments. One of the experiments consisted in presenting two stimuli in close succession to participants. Participants were instructed to indicate whether the stimuli were identical (*Same*) or if one element differed (*Different*). After inspecting the data in the middle of data collection (which you should never do![2]), Nickerson noticed a

---

[1] Even though the article was published in 1965 (see reference for Nickerson, 1965), it was accepted in October 1964.

[2] The inspection of data in such a way can lead to an inflation of the reported effect, or *data dredging* (Brodeur et al., 2012; G. D. Smith & Ebrahim, 2002). It is now generally seen as a bad research practice because it adds a supplementary degree of freedom (the researcher's desire for positive results). One of the consequences of such practice is that results are often difficult to replicate, or greatly overestimated

strange result: *Same* responses were faster than *Different* responses. Humbly, he provided no explanation as to why this occurred; he just reported it.

A few months later, Bindra, Williams and Wise (1965) published results from a tone matching study, reporting that participants took longer to identify *Same* tones than *Different* tones, thus contradicting Nickerson's finding, who used visual stimuli (letters). This led to a first finding regarding *Same-Different* judgments: the fast-same effect is only observed if the stimuli are codable (Bindra et al., 1968; Grill, 1971; Nickerson, 1969, 1972). In other words, humans manifest a fast-same effect only if they can create a mental representation of the stimulus.

However, it did not explain why the fast-same effect occurred for codable stimuli. Nickerson (1967b, 1967a, 1968) wrote three more articles, essentially showing that the effect is bizarre, and then wrote two book chapters (Nickerson, 1973b, 1978) in which he reviewed various studies on the fast-same effect. Nickerson did not solve the fast-same problem, but he showed to the scientific community that it was an effect worth investigating. After over a decade of research on the topic, he left the field but let behind him potential theoretical explanations. He suggested exploring the potential effects of priming, attention and inherent preference for similarity on response time in *Same-Different* tasks.

(Ioannidis, 2005, 2012; Open Science Collaboration, 2015; Savalei & Dunn, 2015; Simmons et al., 2011). However, Nickerson reported a fast-same effect of about 45 ms, which does not seem to be an overestimation, speaking from my personal experience.

The period between Nickerson's studies and the mid-1980s may be considered the golden age of *Same-Different* research. It is a tale of many proposed models, unverified assumptions, fights between researchers and, like in any good drama, a tragic ending that leaves the readers without a clear answer. In the first chapter of my thesis, I tell this story of how the scientific community explains that humans are faster to respond *Same*.

Obviously, when I started studying *Same-Different* judgments, I was not aware of this story. My initiation to the fast-same effect was in my honour's thesis project, in which I compared *Same-Different* judgments in letter and face stimuli. The idea was to compare global and local information processing in a *Same-Different* setting. Unfortunately, these data were never published: They were too messy and not conclusive. Nevertheless, I observed the fast-same effect using both stimuli. I also did a few other (unpublished) experiments, using words and non-words stimuli, or varying the inter-stimuli interval. Those were my first experiences with the frustrations of inconclusive results and rejections for publication. I did publish one article on *Same-Different* data, in which I fitted sequential sampling models to three *Same-Different* tasks (Goulet & Cousineau, 2019a). I also had the chance to work with many colleagues who also studies *Same-Different* judgments, some of which had more success with publications (T.-Groulx et al., 2019; Walker & Cousineau, 2019).

After many years and many projects on this topic, I found myself just seeking for the answer regarding how humans are faster to respond *Same*. I do not prefer an answer over another. The *Same-Different* research, for some reason, stopped in the mid-1980s without a satisfying answer. The most recent review article on the fast-same effect is Sternberg (1998), which barely cites any studies conducted in the 1990s.

Since then, research in cognitive psychology has changed a lot. With better access to computational power, researchers now have better tools to interpret data and answer their research questions. In the first study of this thesis, I analyze the data using the *Linear Ballistic Accumulator* model, developed by Brown and Heathcote (2008). In the second and third studies, I use double-factorial paradigms and *Systems Factorial Technology*, developed by Townsend and Nozawa (1995), to assess the cognitive architecture and the processing capacity underlying *Same-Different* judgments. I have the chance to benefit from these recent advancements in modelling and statistical tools to delve again in the fast-same effect. I do not have the pretension of saying that I solved the problem, but I think I was able to narrow it down to a few answers.

This thesis contains six chapters. The first chapter traces the history of the *Same-Different* research and summarizes the many models proposed to explain the fast-same effect. The second chapter provides an overview of the three experiments I conducted to test these various models. Chapters three is the manuscript accepted for publication by the *Journal of Experimental Psychology: Human Perception and Performance*. Chapter four is the manuscript submitted *Acta Psychologica* for publication. Chapter five is the manuscript accepted for publication by *Attention, Perception, & Psychophysics*. Because they are standalone papers, the introductions of these articles will be redundant with the first and second chapters of this thesis. Nevertheless, they each focus on questions specific to the article's main research question. In the sixth chapter, I synthesize the main findings of the three studies and answer the main question of this thesis. Finally, I end with personal considerations, discussing the implications of my research and the methodological particularities of this thesis.

Chapter 1 – Model Farm

The fast-same effect is the tendency for participants to respond, on average, *Same* faster than *Different*, a speed advantage of about 35 to 50 ms (Bamber, 1969; Nickerson, 1965; Posner & Mitchell, 1967). Most studies that reported the effect used letters and digits (e.g., Bamber, 1969; Silverman & Goldberg, 1975; Snodgrass, 1972), but some used more unorthodox stimuli, such as geometrical shapes, dot patterns, colours and motion direction (e.g., Dyer, 1973; Egeth, 1966; Hock, 1973; Petrov, 2009). Probably the main reason letter stimuli are so often used is because there are easy to manipulate. Indeed, researchers could vary the complexity of the stimulus (changing the number of letters in a string), the number of mismatching elements (e.g., one mismatch out of four letters compared to three mismatches out of four letters) and the level of identity (i.e., physical matching, nominal matching and categorical matching).

The first major *Same-Different* study was the one of Bamber (1969), in which he presented participants with two letter strings sequentially. The task of the participants was to indicate if all the letters in the string were matching or if at least one of them was mismatching. He showed that increasing the number of letters in a string also increases the response latency in both *Same* and *Different* trials. He also showed that in *Different* trials, decreasing the number of mismatching letters also increase response times. In other experiments, Bamber asked participants to match uppercase letters to their lowercase counterparts (i.e., “J” and “j” are *Same*; “J” and “c” are *Different* – Bamber, 1972; Bamber & Paine, 1973). The idea of examining various levels of comparison originated from the work of Posner and Mitchell (1967). These authors found that having more abstract level of comparison diminishes the fast-same effect, but that *Same* responses have nonetheless faster response times. This finding was supported by many subsequent studies (Beller, 1970; Pachella & Miller, 1976; Proctor, 1981, 1986; Well & Green, 1972).

These results share some similarity to those of visual search tasks, which were also a topic of interest in the 1960s. In such tasks, participants must search a target among many distractors. One could argue that it is almost mirroring the *Same-Different* task: instead of searching for mismatching elements, participants are searching for matching elements. A result often associated with the visual search tasks is that as the number of distractor increases, participants tend to take more time to identify the target (Atkinson et al., 1969). However, in visual search tasks, the position of the target and distractors differ between trials. In *Same-Different* tasks, the stimuli are displayed at constant locations; thus, participants are truly only looking at one stimulus and its composing elements.

To better compare the two settings, researchers concurrently developed what is now referred to as the *disjunctive* task (Bamber et al., 1975; Derks, 1972; Farell, 1977; Nickerson, 1967a; Sekuler & Abrams, 1968; Silverman & Goldberg, 1975; D. A. Taylor, 1976a). To avoid confusion, some researchers therefore refer to the *Same-Different* task as the *conjunctive* task. The two tasks are identical, except that in the (more usual) conjunctive task, participants must respond *Same* only when the stimuli are in complete conjunction (when all the elements are matching), whereas in the disjunctive task, participant must respond *Different* only when the stimuli are in complete disjunction (when all the elements are mismatching). For example, in the conjunctive task, the stimuli “JCVD” and “JCGP” is considered *Different* (the last two letters are mismatching), but in the disjunctive task, it is considered *Same* (the first two letters are matching).

Another way to conceptualize these two tasks is by referring to the concept of *critical elements*. In the conjunctive task, the critical element is a mismatching letter. To respond *Different*, participants only need to identify one mismatching letter (the critical element of the task). Conversely, in the disjunctive task, to respond *Same*, participants only need to identify one matching letter (again, the critical element of the task). The two tasks only differ on the identity of the critical elements. Therefore, the effect of the number of critical elements should be replicated in each task.

In the disjunctive task, the response time is inversely proportionate to the number of critical elements: the more elements that are matching, the faster is the response time. This means that the slowest condition is the one without any critical dimension (the *Different* response). It is logical: participants must exhaust that all the elements are mismatching to respond *Different*. Following this logic, we would expect to find an identical effect in the conjunctive task. We do find a similar inversely proportionate relationship between the number of critical elements and response times, except for one condition, when the correct response is *Same*. In this condition, there are no critical elements and response times should be the slowest. However, as it has been mentioned many times now, we find quite the opposite: this is the fastest condition. This is why the fast-same effect is bizarre: in the conjunctive task, participants need to exhaust all the matching elements of the stimulus before responding *Same* but identify one mismatching element to respond *Different*.

This led researchers to believe that there is probably something unique about the conjunctive task, something that disappears in a disjunctive setting. Participants might be able to accumulate information about the absence of elements, just like they can accumulate information about the presence of an element (an argument recently supported by electroencephalographic data; Hyun, Woodman, Vogel, Hollingworth, & Luck, 2009). However, there might be structural differences between these accumulation processes that explain the fast-same effect. In the remainder of this chapter, I review the structural differences that were proposed by various researchers. Many have done a similar work in the past, running studies showing that there exists a plethora of models, which all seem to explain why the fast-same effect occurs (Farell, 1985; Krueger, 1978; Nickerson, 1973b, 1978; Ratcliff, 1978, 1981, 1985; Sternberg, 1998). The key word here is *plethora*, meaning not only a large number, but an excessive number of explanations, some of which are contradicting.

The Tortoise and the Hare

The first concrete model aiming to explain the fast-same effect was formulated by Bamber (1969). He originally hypothesized that participants use a serial self-terminating strategy to perform the task in which they attend to each letter in succession until they find a mismatch. He found that this model fits *Different* conditions well, but that it poorly fits the *Same* conditions. Bamber therefore suggested that participants use two processes to realize the task. In this dual-process approach, participants first process all the letters in parallel, as a *chunked* piece of information, akin to template matching. If this fast processor (also called an *identity reporter*) fails to respond *Same*, it cedes its place to a more analytical processor (which is serial). The model seems congruent with the proposition of Garner and Clement (1963) that participants can process stimuli at a global level, but if needed, can focus on its composing elements (also see Angiolillo-Bent & Rips, 1982).

There were many supporters of the dual-process model (Decker, 1974; Derks, 1972; Krueger, 1973; Millspaugh, 1978; Nickerson & Pew, 1973; Silverman & Goldberg, 1975; D. A. Taylor, 1976b; Tversky, 1969). The solution is quite simple and seems to fit with the popular idea that humans are good at chunking information. That said, Bamber was not entirely convinced. In his original conception, the identity reporter is limited to physical matching and would not extend to situations in which participants would compare stimuli based on nominal information (i.e., matching an uppercase letter to its lowercase counterpart). Similar to the work of Posner and Mitchell (1967), and Beller (1970) on nominal identities, he argued that his model was necessarily erroneous because *Same* trials were still faster even when the two stimuli are not perfectly matching (Bamber, 1972; Bamber & Paine, 1973). The effect was reduced (Well & Green, 1972), but it was present nevertheless.

A possibility is that participants do not process solely physical information with the identity reporter, but also related information, such as auditory similarities (*phonological* information), or categories (i.e., vowels and consonants). This means that participants could rely on multiple sources of information to respond *Same* (B. A. Eriksen & Eriksen, 1974; Posner & Mitchell, 1967). Additionally, it means that participants would be somewhat flexible with what they considered a *Same* stimulus. This assumption was tested in the 1980s by researchers who observed that the fast-same effect still occurs even when the stimuli were rotated (Bagnara, Simion, & Umiltá, 1984; Larsen, & Farrell, 1981; Koriat & Norman, 1988, 1989a; Simion, Bagnara, Roncato, & Umilta, 1982). The effect was modulated by the size of the rotation, that is longer response times when the stimuli rotation was greater, similar to what is reported in the seminal paper of Shepard and Metzler (1971). Perhaps the most important finding in these studies was that only the response times in the *Same* condition were affected by the rotation, which had no effect on *Different* response times (Koriat et al., 1991). This further reinforces the idea that the judgment of sameness and differentness are part of two separate processes.

The works on nominal identity and mental rotation suggest that participants must be able to accurately represent the stimulus mentally to respond *Same* faster. A failure to code the stimulus and its labels results in a failure to benefit from matching identity. This would explain why in experiments that used non-codable stimuli (i.e., sound tones, line length), participants are slower to respond *Same* compared to *Different* (Bindra et al., 1968, 1965; Nickerson, 1969). When the stimuli are not codable, participants seem to use a more serial and analytical processing, suggesting the absence of an identity reporter (Grill, 1971).

According to the dual-process model, participants respond *Different* using this analytical processor. Nickerson (1965) suggested that participants might need to process *Different* stimuli more thoroughly. Indeed, whereas participants have a clear mental representation of a *Same* stimulus (and therefore can use template matching to respond *Same*), they cannot generate a clear mental representation of a *Different* stimulus. Hence, participants need a processor capable of more detail-oriented treatment of the stimulus. This view was supported by many researchers (Beller, 1970; Decker, 1974; Derks, 1972; Farell, 1977, 1988; Grill, 1971; Hock, 1973; D. A. Taylor, 1976b; Tversky, 1969) and is congruent with the pioneering works of Sternberg (1966), who proposed that humans usually process information serially.

Howell and Stockdale (1975) tested the plausibility of this serial analytical processor in the *Same-Different* task by comparing participants' sensitivity to the position of the difference in the stimulus. They found that participants tend to process information at the extremities of the stimulus first, then information in the middle. This tendency is observed even when the probability of a mismatching element being in the middle of the stimulus is higher than at the extremities (Proctor et al., 1991).

Even though most researchers supported the serial view of the analytical processor, a few suggested that it might process the information in parallel (Donderi & Zelnicker, 1969; Hawkins, 1969; Hawkins & Shigley, 1972). They proposed that perhaps participants are processing all the information simultaneously, but that some elements (for instance the ones at the extremities) have a faster competition time. However, this view had few supporters and has never been supported directly (that said, neither has the serial view).

Currently, the cognitive architecture underlying *Same* and *Different* judgments is only assumed. So is the capacity of the processing system, a question that has been vastly ignored in the *Same-Different* literature. This is surprising considering the fact that Hylan, as early as in 1903, discussed the importance of processing capacity in the understanding of fundamental cognitive abilities. Processing capacity refers to the efficiency at which participants process individual pieces of information. This supposes that participants process single element of information at a given rate. However, this rate can potentially be affected by the presence of multiple other elements that need to be processed.

Almost all researchers assumed that this was not the case and that the number of elements in the stimulus does not affect the processing capacity of the cognitive system. It is unclear whether these researchers voluntarily assumed *unlimited* capacity or simply ignored this aspect. A good analogy to capacity is the *pool of resources*. In order to process a stimulus, participants need some resources. If there is an unlimited amount of resources, participants would process single elements at the same rate, regardless of the total number of elements contained in the stimulus or any other condition which would increase the cognitive workload (e.g., time pressure). Basically, increasing the workload means that more resources are needed to accomplish the task – that is to process the stimulus. If the resources are limited, participants would be less efficient under high workload conditions than under low workload conditions; their processing capacity would be *limited* (Link & Tindall, 1971).

Another possibility is that the presence of multiple elements could act as a form of redundancy in the signal and enhance the processing capacity of a cognitive system (J. Miller, 1978, 1982). Miller viewed the cognitive system as combinations of multiple channels responsible to process certain elements of a stimulus. For example, in a *Same-Different* task with letter strings, a four-letter stimulus would require four channels. Independently one from another, the channels accumulate information about the response alternatives. Say three of those four letters are mismatching, it would mean that there exists redundancy in the accumulation process, which ends up increasing the efficiency of the cognitive system. Under these conditions, the system is working in *super*-capacity.

Much like limited capacity, there is not a lot of support in the *Same-Different* literature for super-capacity. However, there is some support for the redundancy effect (Holmgren et al., 1974; J. Miller, 1982; Proctor & Healy, 1985; van der Heijden et al., 1984). Capacity was not explicitly part of many discussions; researchers did not directly address this cognitive aspect of the task. It also did not fit well with the dominant serial self-terminating view of most analytical processors. Such processors imply that one channel can respond independently of the others. For example, if a mismatching letter is detected at the second processed letter, the treatment can be ceased by the channel responsible to process the second letter. However, super-capacity implies that the processing of other channels somehow increases the efficiency of the whole system.

One possibility is that there exists a cognitive unit dedicated to generating a response and that the individual processing channels are simply feeding their accumulation to that unit. If such unit exists, redundancy in the signal would mean that it reaches its decision threshold faster; that it is working in super-capacity. This architecture is neither serial nor parallel, but coactive. Another possibility is that the processing channels can interact (also known as *cross-talking*). This would mean that one channel accumulating mismatching information would further activate the *Different* signal and inhibit the *Same* signal in other channels. This theory was supported in a study that used words and non-words as stimuli (Krueger & Shapiro, 1979). They found that participants were faster to respond when the stimuli were words compared to non-words. This is likely because in words trials, participants inferred some letters of the stimulus while processing the other letters (they could accurately *guess* a letter without processing it). They could not do such inference in non-word stimuli.

Slowly, researchers started to consider alternatives to the dual-process model. Due to the lack of direct testing for cognitive architecture and little consideration for the role of processing capacity, it is possible that *Same-Different* judgment is not the fruit of two processors: a rapid hare to respond *Same* and a slower but steadier tortoise to respond *Different*. If the analytical processor is affected by workload (either negatively or positively), why would the identity reporter also not be affected by it? Finally, if the identity reporter fails to identify that the two stimuli are identical, why would participants need a more thorough treatment of the stimulus to respond *Different*? After all, it is the only other response alternative. Under such severe accusations, the dual-process model was almost abandoned by researchers in the 1980s, leaving its place to more parsimonious single-process models.

The Improbable Duo of Red Pollard and Seabiscuit

In the *Same-Different* task, participants do not need to indicate the number of mismatching elements, nor do they need to locate them. They simply need to respond *Same* or *Different* as fast and as accurately as possible. The dual-process model portrays this race as one between a tortoise and a hare. Sure, in the fable, the tortoise ends up winning the race, but nobody doubts that if the hare was seriously racing, it would be the fastest. It could probably win even before the tortoise started to move. Single-process models propose a fairer race, almost to the photo-finish, more similar to a race between horses. After all, in a horse race, a difference of 40 ms would probably require a photo-finish. That said, a fast horse is a fast horse and it will regularly finish faster than its competitors.

Perhaps responding *Same* is like riding Seabiscuit, the world-famous racing horse. Even if the name Seabiscuit is now synonymous with great speed, its success is quite mysterious. The horse was, according to many jockeys, uncontrollable. Nobody could mount it except a tall and mediocre Canadian jockey named Red Pollard.[3] Seabiscuit and Red Pollard are just like *Same* responses: they should not win the race; they should be last. But somehow, their duo is fast, and many have their theories as to why this occurs. In this section, I describe these many theories (about the fast-same effect, not about Seabiscuit), in their historical order.

Detecting Mismatches in the Presence of Noise

Perhaps the biggest critique of the dual-process model is its inability to explain why participants need to process *mismatches* at all, since a failed detection by the identity reporter means that the correct response is necessarily *Different*. Before Bamber (1969), a few authors proposed that *Different* responses might require a more thorough analysis of the stimulus (Egeth, 1966; Nickerson, 1965).

In visual search tasks, participants are known to come back to some stimuli that they already processed, a phenomenon known as *rechecking* (Howell & Stockdale, 1975). Re-checking a stimulus increases the total response time at the benefit of increasing confidence in the response (and its accuracy). Krueger (1978) suggested that participants

[3] Red Pollard might have been the same height as Napoléon Bonaparte (1.70 m), he was noticeably taller and heavier than most competitive jockeys, putting him at a great disadvantage. In horse racing, a small and light jockey means that the horse has less weight to pull and therefore should be faster.

committed more rechecks in *Different* trials compared to *Same* trials, which would result in the fast-same effect.

Krueger's formulation was highly inspired by the famous decision-making *Signal Detection Theory* (SDT; Tanner & Swets, 1954). According to SDT, participants ponder response alternatives by encoding stimuli with a certain efficiency and by comparing this signal to a decision criterion. Krueger argues that the single processor responsible for *Same-Different* judgments – what he calls a *noisy operator* – encodes matches and mismatches with the same efficiency (Krueger & Shapiro, 1981). He also proposed that the *Different* response criterion is higher than the *Same* response criterion; that participants need a stronger signal to respond *Different* than *Same*.

The rationale behind this *criterion shift* is that the perception of a stimulus is noisy. In other words, the letter "J" might not be perceived as a "J" first, it might look like an "I" or a "T". To be sure that the letter is indeed a "J", participants must recheck the stimulus numerous times. However, participants already know what a *Same* stimulus looks like. They also know that it is more likely that they perceive a matching letter as a mismatching letter (due to visual noise) than a mismatching letter as a matching letter. For example, the first stimulus could be "JC" and the second stimulus "JD". After a few rechecks, a participant identifies that the first letter of the second stimulus is a "J", which is likely, as it was also the first letter of the first stimulus. They also identify that the second letter of the second stimulus is a "D", but they also need to consider the possibility that it might also be a "C" that looks like a "D" due to noise. To rectify this uncertainty, the participant rechecks the second letter a few more times, resulting in a slower response time.

Contrary to the dual-process model, the noisy operator model argues that there is nothing special or unique about the processing of *Same* information. The single processor accumulates information about matches and mismatches in the same fashion, but the response criterion for *Same* is lower than the one for *Different*. Supporting this model, Crist (1981) showed that participants were also faster to respond *Different* when the two letters were dissimilar as opposed to similar. For example, participants were faster to respond *Different* when comparing a “J” to a “V” than when comparing a “J” to a “T”.

An alternative approach would be to remove noise from the model and replace it with inhibition. When the two stimuli are presented successively, the encoding of mismatching letters in the second stimulus is inhibited because of the presence in working memory of the letters of the first stimulus. However, the inhibition is considerably reduced in simultaneous presentation. Likewise, because information in working memory tend to decay (Di Lollo, 1980), increasing the temporal separation between the presentation of the two stimuli also reduces the fast-same effect. This temporal modulation of the fast-same effect is not predicted by the noisy operator model, and Krueger conceded that his model was probably erroneous (Chignell & Krueger, 1984).

Repetition Effect

What most contributed to the rejection of the noisy operator model was the *sensitivity shift* model of Proctor (1981). The model is in many aspects like the noisy operator: it has a single processor and it is highly inspired by SDT. The major difference is that instead of a discrepancy between decision criteria, *Same-Different* judgments are characterized by distinct encoding sensitivity (Krueger & Shapiro, 1981; Proctor & Rao, 1983a). According to this model, participants can encode matching information faster than mismatching information.

This occurs for two reasons. First, the presence of some letters in working memory *facilitates* the encoding of the matching letters in the second stimulus. Consequently, encoding matches is faster. Second, the presence of these letters also *inhibits* the encoding of mismatching letters in the second stimulus, which slows down the encoding of mismatches. Because there are two effects occurring at once (facilitation and inhibition), the model can explain why the fast-same effect is reduced in simultaneous presentation settings. It can also explain why the fast-same effect is reduced when the level of comparison changes (e.g., nominal task vs. physical task): there is no facilitation due to repetition, but inhibition still occurs (Eviatar et al., 1994).

The idea that the repetition of the stimulus could explain the fast-same effect was not novel (Entus & Bindra, 1970; D. A. Taylor, 1977). It is highly probable that participants are primed to respond *Same* just because they already saw the stimulus. The idea that the encoding speed of a single processor can be affected by the experimental design or the level of comparison was not novel either (J. O. Miller & Pachella, 1973; Pachella & Miller, 1976; Well et al., 1975). Proctor’s sensitivity shift model was able to merge those propositions together into a formalized model. That said, this model is portrayed as a one-time effect of sensitivity and not much as a dynamic process. This is probably because SDT is a probabilistic model that aims to predict response proportions, not response times. Under the SDT framework, the encoding sensitivity rate does not vary over time, neither does the decision criterion. In other words, the model is not stochastic.

Response Competition

The lack of consideration for the encoding rate is probably the main critique of the sensitivity shift model. Because of this, some authors argued that inhibition is too pervasive in the model (C. W. Eriksen et al., 1982; St. James & Eriksen, 1992). Instead, these authors proposed to extend the sensitivity shift by adding response competition, a temporal element that can both increase and decrease the encoding speed of information. The response competition model adds only one more assumption: that the processing channels can influence each other – that they crosstalk (Krueger, 1987; Pan & Eriksen, 1993; D. A. Taylor, 1977).

For example, consider a trial in which the first stimulus is “JCVD” and the second stimulus is “JCGP”. Here, the correct response is *Different*, but the first two letters are matching. If channels crosstalk, the channels responsible to process the first and second letters would inhibit the encoding speed of the channels responsible to process the third and fourth letters. This is because encoding information in favour of the *Same* response creates competition to the *Different* response. However, in *Same* trials, all the letters are matching, and the response competition is almost (if not completely) absent.

Despite this innovation, the response competition model was not very popular. There are a few reasons why this was the case. First, it fails to explain, alone, why *Same* responses are still faster than all-*Different* trials (i.e., trials with only mismatching letters). In such trials, there is no response competition, but the *Different* trials are not that much faster; they are still aligned linearly with the other *Different* conditions that contain one or more matches. Second, like the noisy operator model, it fails to explain how very dissimilar letters can create response competition (Nickerson, 1981). The only way the response competition model would work is if crosstalk between the processing channels speeds up the encoding of *Same* information more than the encoding of *Different* information. However, such asymmetry is unlikely and is not a strong argument against the sensitivity shift model. Finally, the response competition model probably went under the radar because of another model that was in opposition to the sensitivity shift model: the response bias model.

Response Bias for Sameness

The criterion shift and sensitivity shift models are not incompatible, but they explain the fast-same effect using too many mechanisms. After a few articles comparing the two models (Chignell & Krueger, 1984; Krueger & Shapiro, 1981; Proctor & Rao, 1983a), Krueger conceded that the sensitivity shift model is more generalizable to various manipulations of the *Same-Different* task. As for the response competition model, it did not attract much attention, due to its limitations, but also due to the predominant place that took the response bias model.

Ratcliff (1978) developed a broad model of memory retrieval he called the *drift-diffusion model*. The model is stochastic, in the sense that it describes the accumulation of information over time. It shares some characteristics with Krueger's (1978) conceptualization: a single processor that accumulates information about response alternatives at a certain rate (sensitivity) until the signal is strong enough to reach one of the decision criteria. That said, the drift-diffusion model was more complex, adding some supplementary parameters, including within- and between-trial accumulation rate variability, extra time for generating a motor response, and response bias.

Ratcliff (1981) applied the drift-diffusion model to the *Same-Different* task and argued that the fast-same effect was not a question of sensitivity or criterion (which he calls *accumulation rate* and *decision threshold*), rather a matter of response bias. Simply, humans have an inherent preference for positive responses, including for similarity (an idea also proposed by Taylor, 1977). This inherent bias could be the result of evolutionary processes which result in faster response times for *Same* responses compared to *Different* responses.

The response bias model predicts that one way to disrupt the fast-same effect is to create an artificial bias in favour of the *Different* response. Indeed, when participants are instructed to be cautious when they respond *Same*, the fast-same effect disappears (Ratcliff & Hacker, 1981). The authors argued that this effect is not predicted by the sensitivity shift model, as repetition effect should nevertheless be observed under such instructions. Proctor and Rao (1982) replied that even though the fast-same effect was disrupted, it was not completely reversed, as Ratcliff and Hacker originally suggested. When participants are instructed to be cautious when they respond *Same*, the fast-different effect is not as large as the fast-same effect typically observed in classical settings. The difference between the fast-different and the fast-same effects was around 50 ms. Therefore, Proctor and Rao concluded that the sensitivity shift due to repetition is still present, even when participants are instructed to be cautious about *Same* responses.

This reply lead to multiple exchanges of arguments in favour and against the two models (Proctor, 1986; Proctor & Rao, 1983b; Proctor et al., 1984; Ratcliff, 1985; Ratcliff & Hacker, 1982; Ratcliff et al., 1989). Both clans disagreed over the interpretation of empirical results, especially in the task with strong imbalance in responses (i.e., 90% of the trials are *Different* and 10% are *Same*, and vice versa). A decade earlier, Nickerson (1973a) explored the effect of the frequency of response, but reported that both *Same* and *Different* stimuli were affected by frequency in a similar fashion, and consequently did not continue research on this specific topic.

Ratcliff (1985) fitted the drift-diffusion model to the data of Proctor and Rao (1983b) manually and showed that he could fit *Same-Different* results very well by varying the bias parameter. Proctor criticized the model fit, arguing that its good fit was "probably a consequence of the simulation being generated without a theoretical rationale to constrain parameter values, rather than being due to a property of the human information processing system" (Proctor, 1986, p.476).[4]

This war of models led to the end of the *Same-Different* research golden age. It ended with a lot of heated arguments in favour and against two incongruent explanations for the fast-same effect. The sensitivity shift model argues that it is caused by a repetition effect, therefore is stimulus-induced. The response bias model argues that it is caused by an inherent preference for *Same* response, therefore is participant-induced.

Attention as a Modulator of Information Processing

In the shadow of these arguments, a final model emerged, first developed by Farell (1977), in his doctoral thesis. Farell was keen on linking the conjunctive and disjunctive tasks and provided a theoretical model that can account for both tasks. He was inspired by the dual-process model, especially by the identity reporter, which was likely

---

[4] In the recent years, there were multiple discussions on model fitting, constraints decision and theoretical support of parameter values, which supports Proctor's claims (Navarro, 2019; Roberts & Pashler, 2000).

pre-attentive (Beller, 1970). However, Farell did not conceptualize the identity reporter as a distinct processor, but as a processing stage of the unique processor responsible for both *Same* and *Different* responses. In other words, there is only one processor, but it can work in multiple fashions – it has different stages.

The first stage, associated with *Same* responses, consists in mapping the first stimulus to the only possible identical second stimulus, what Farell calls *one-to-one mapping*. For example, in the conjunctive task, if the first stimulus is "JCVD", there is only one possible second stimulus that would trigger a *Same* response: "JCVD". According to this model, the processor does not need attention to perform such task. On *Different* trials, however, a *Different* stimulus can be anything. Participants must then perform a *many-to-one mapping*: they must consider many alternatives simultaneously. Likewise, in the disjunctive task, participants also need to perform a many-to-one mapping for *Different* trials (the condition with no critical element). Such mapping is likely too demanding to the participants (they likely do not keep all possible *Different* stimuli in working memory). Therefore, the processor requires attention to analyze the second stimulus and compare it to the first stimulus.

In a nutshell, the *attention-driven single processor model* uses attention as a proxy to determine the stage of the processor (Farell, 1977, 1984, 1985, 1988). This model states that attention is demanding for participants and that if they can, they will prioritize not using it. However, if the pre-attention one-to-one mapping is inconclusive, participants will need to use attentional resources to respond, which takes more time, but also likely affects processing capacity negatively. This model is congruent with the theory of pattern recognition (Neisser, 1967), where objects kept in memory are a chunked mental representation of their many elements and put in categories (Buffart & Geissler, 1984). After Farell (1988)'s paper, research on *Same-Different* judgments has slowed down, and the attention-driven one-process model was added to the list of unrefuted models.

Modern Times

At the turn of the millennium, researchers attempted to investigate *Same-Different* judgments from a different perspective, namely neuroscience. Perhaps some brain areas are associated with such judgments and that it could reveal something about the cognitive processes at play. The idea to relate brain areas to *Same-Different* judgments was not completely novel. Morais and Darwin (1974) presented auditory stimuli in a monaural fashion and found that there is a hemispheric preference for *Different* trials, but not *Same* trials. Indeed, whereas response times were almost identical in *Same* conditions, participants were on average 16 ms faster to respond *Different* when the stimulus was presented in their right ear compared to the left ear. They interpreted this result as an evidence that *Different* judgment requires a more analytical processing than *Same* judgments, and that such processing is specific to the left hemisphere.[5] This study was seemingly unnoticed by other *Same-Different* researchers and only attracted the attention of a few authors in language processing.

Hübner (1998) ran a similar study, using visual stimuli instead, namely Navon letters (see Navon, 1977), and focusing on the hemispheric specificity of global and local

---

[5] Auditory percepts are contralateral, meaning that a stimulus heard with the right ear is sent to the left hemisphere first, and vice versa.

information. Contrary to Morais and Darwin (1974), he found no hemispheric differences in *Same* nor in *Different* trials. Naikar (1999) reported that patients with a commissurotomy were unable to perform *Same-Different* judgments when visual stimuli (i.e., colour and motion direction) are presented to both the left and right visual field (to both hemispheres), but that they could do the task with accuracy when the stimuli are presented only to a specific hemisphere. This suggests that *Same-Different* judgments are not restricted to one hemisphere. That said, fMRI data suggests that there might be a brain area (regardless of the hemisphere) that correlates specifically with *Different* judgments: the caudate (Sinha & Glass, 2017). These authors interpreted this result as evidence in favour of the dual-process model.

These results seem to converge towards the idea that the detection of *Same* information is automatic and that in order to respond *Different*, participant needs to process the stimulus further. However, this view is not unanimous, and other researchers proposed that change detection is the default mode of processing, not similarity detection (Lachmann, 2001; Lachmann & Geissler, 2002). This was inspired by a popular new model of memory retrieval, the exemplar-based random-walk (Cohen & Nosofsky, 2000; Nosofsky & Palmeri, 1997). According to this model, participants can, by default, process and detect changes, but also actively search in their working memory for matching features. This active memory search result in faster accumulation of information in *Same* trials, whereas the more passive detection of change is characterized by slower accumulation rate. Some authors found evidence that participants are indeed able to detect changes (Davelaar et al., 2011; Hyun et al., 2009), even in olfactory stimuli (Møller et al., 2012).

These recent studies seem to support the idea that the fast-same effect is likely the result of the facilitation of the *Same* decision compared to the *Different* decision, and not of response bias. Some authors highlighted that the latter explanation is not found when fitting other models to the data than the drift-diffusion model (Goulet & Cousineau, 2019a; Van Zandt et al., 2000). That said, the response bias model still has defenders and should be considered as valid potential explanation for the fast-same effect (DeCarlo, 2013; Irwin et al., 2001; Ratcliff & McKoon, 2008).

As for the sensitivity shift model (or the decision facilitation explanation), the question remains: what are the sources of that facilitation? When two identical visual stimuli are presented, they do not only match on their physicality, but also nominally, phonetically, categorically and semantically. Potentially, all these sources of activation can contribute to the facilitation of the *Same* response accumulation (Belke & Meyer, 2002; Sagi et al., 2012), although the role of phonetic information remains unclear (Kinoshita & Kaplan, 2008; Lupker et al., 2015; Walker & Cousineau, 2019). Fluctuation in attentional resources might also modulate response facilitation (Jacob et al., 2013). The redundancy in the signal for the *Same* response could explain why the accumulation rate is faster than for *Different* response (Goulet & Cousineau, 2019a; Harding, 2018).

Chapter 2 – Presentation of the Thesis's Studies

The literature on the *Same-Different* task is quite rich for such a simple task. From the participants' perspective, the task is so easy that they seem to complete it without any effort. They often report that they rapidly enter a *mental zone* in which they feel like they shut down their brain and just do the task. In principle, this should be easy to study: There is little exterior noise and participants do not overthink the task, they just do it. However, after over 50 years of research, there is no consensus on what explains the main result of such task, the fast-same effect.

I think the reason why this is the case is because most researchers conducted studies with their own model of the task in mind. With a clear model of the task, these researchers were able to test specific questions and test hypotheses deduced from their model. If a hypothesis is confirmed, it serves as evidence for the model. For example, Ratcliff (1985) fitted the drift-diffusion model to Same-Different data and hypothesized that the bias parameter would vary as a function of the response, which is what he found. Later, Cohen and Nosofsky (2000) fitted the exemplar-based random-walk model to Same-Different data and hypothesized that the accumulation rate parameter would vary as a function of the response, which is what they found. Technically, both these studies show support for alternative models, which makes them more plausible. That said, confirming hypotheses is not conclusive evidence for a model , at least from the perspective of falsificationists (Popper, 1959). And this is clearly shown with these two studies. Taken individually, they both seem to have found the solution to the fast-same effect, but neither model can coexist: one has to be false – and likely both are (Navarro, 2019).

Perhaps a more judicious approach is not to enter a quest of finding evidence in favour of a model, but instead to seek for data that refute the models. For example, Chignell and Krueger (1984) tested a hypothesis made by the criterion shift model that spatial separation between stimulus should affect response time more than temporal separation. The results suggested the opposite: that temporal separation affected response time and that spatial separation did not. The authors conceded that the criterion shift model was probably not a valid explanation to the fast-same effect and that alternative models seemed more adequate, namely the sensitivity shift and the response competition models. Model falsification, especially when one of the authors is the instigator of the model itself as in the above example, is rare. This is probably because most models are tested using methodologies from which they were developed (Kuhn, 1961). Innovation is key to test models: new tasks must be created to test specific hypotheses for which the models have a legitimate chance of being falsified.

I approached the *Same-Different* task with this mentality. I wanted to test specific hypotheses in a creative way to put the models to test. Therefore, I developed three *Same-Different* task variations that were never done and focused on precise research questions that are central to the proposed models of the task and the fast-same effect. In this chapter, I present these studies, but more importantly, the precise research questions and hypotheses that I intended to test.

Study 1 – The Participant or the Stimulus

The two models that had the most impact on the recent *Same-Different* literature are probably the sensitivity shift (which to ease a comprehension, I will henceforth call the *facilitation model*) and the response bias model. These two models were in clear opposition. Facilitation is seen as a way to enhance the treatment of a stimulus and is much in line with the concept of priming, a phenomenon whereby the treatment of a stimulus is augmented by prior information. According to the facilitation model, the processing of the second stimulus is enhanced by the encoding of the first stimulus; it is a repetition effect. This enhancement can occur because of various sources (e.g., physical, phonological, semantic), but overall, it attributes the fast-same effect to faster treatment because of previously presented information. In other words, it is caused by the stimulus.

The response bias explanation also has many supporters, mostly because of the success of the drift-diffusion model to fit various cognitive tasks (Ratcliff & McKoon, 2008; P. L. Smith & Ratcliff, 2015). When fitted to Same-Different data, the drift-diffusion model suggests that response bias is more involved than the accumulation rate in explaining the fast-same effect – although Goulet and Cousineau (2019) did not replicate this finding.[6] In this model, *Same* responses are faster than *Different* responses because participants have an inherent preference for positive response; the characteristics of the stimulus do not matter.

Identifying the source of the effect is important to better understand the cognitive mechanisms underlying the task. If the fast-same effect is solely caused by participants' bias for the *Same* response, it is assumed that the treatment of the stimulus itself does not

---

[6] One possible explanation for this failure to replicate is that Goulet and Cousineau (2019) fitted the drift-diffusion model using a maximum likelihood technique, which accounts for all response times. In the early days, the drift-diffusion model was fitted manually; more recently, it was fit to quantiles using the Kolmogorov-Smirnoff statistic (Voss & Voss, 2007). Whereas fitting all response times is more thorough, fitting quantiles is more stable, as the model does not need to account for the very fast response time.

change between conditions. However, if the fast-same effect is caused by the characteristics of the stimulus, it supposes that the processing of *Same* and *Different* stimuli is distinct. Not only is this a critical assumption made by the facilitation model, but it is also one made by the dual-process model, and by the attention-driven single-process model. Consequently, I identified this key assumption as the first research question of this thesis: is the fast-same effect caused by the stimulus or by the participant?

To answer this research question, I created a variation of the *Same-Different* task which I call the *Exclusive-Or* task. Instead of responding *Same* or *Different*, participants are instructed to respond *All* if all the elements of the second stimulus are the same or if they are all different compared to the first stimulus, and *Some* if some elements of the second stimulus are the same and some other elements are different. For example, if the first stimulus is “J C”, an *All* stimulus could either be “J C” or “V D” and a *Some* stimulus could be “J D” or “V C”. The objective of this experimental manipulation is to prevent response bias by making participants press the same answer key for both *All-Same* and *All-Different* trials. If participants are biased in favour of a response, they should be regardless of the type of response. In other words, *All-Same* and *All-Different* trials should have comparable response times. However, if the fast-same effect is caused by the characteristics of the stimulus, participants would not be affected by this manipulation and still respond faster in the *All-Same* condition compared to the *All-Different* condition.

In this study, the participants replicated the conditions 200 times. This means that not only the mean response times are quite reliable, but also that the distribution of response times can be estimated in each condition. Therefore, in addition to mean response times analyses, I also fitted accumulator models to the response times distributions. Specifically, I fitted the drift-diffusion model – with little success however – and the linear ballistic accumulator model (Brown & Heathcote, 2008) – with more success. The latter model estimates roughly the same cognitive parameters as the drift-diffusion model, for instance the response bias, the accumulation rate and the motor response time. The major difference between the two models is that the drift-diffusion model states that a single accumulator drifts toward the two response thresholds, whereas the linear ballistic accumulator model states that two accumulators (one for each response alternative) drift toward their respective decision threshold – the first accumulator to reach its threshold triggers the response. There exists more differences between the two models, but they are often used in concurrence because they estimate similar parameters of interests (Donkin et al., 2011; Goulet & Cousineau, 2019a; Heathcote & Hayes, 2012; Ratcliff & Smith, 2004).

The article reporting the results of the first study was accepted for publication by the *Journal of Experimental Psychology: Human Perception & Performance*. Chapter 3 of this thesis is the accepted version of this article.

Study 2 – Architectural Preference for Similarity

A popular idea among researchers is that *Same* information, or at least information related to the similarity of stimuli, is processed in preference over any other type of information. For instance, in the dual-process model, participants rely on the identity reporter to rapidly process a chunked version of the stimulus, and only if this processor fails to respond *Same*, a more in-depth processor is required. In the attention-driven single process model, the default mode of treatment is a pre-attentive one-to-one mapping of the mental representation of the first stimulus and the displayed second stimulus. If this mapping does not conclude in a *Same* response, participants use their attention to process the second stimulus in more details, performing a many-to-one mapping. These models assume that participants process various information about the stimulus in sequence (serially). First, participants inquire about how similar the stimuli are, then they inquire about the other type of information.

The serial assumption is central to explaining the fast-same effect according to these models. However, it is almost ignored by the facilitation and response bias models. Because they are single-process models, they likely assume that all the information about the stimulus is processed simultaneously, but this is not explicitly said. One thing is certain: no one ever directly tested this assumption. Therefore, I decided to address this assumption in the second study of this thesis and focus on the following research question: is similarity processed in preference over other information?

Townsend and Nozawa (1995) developed a statistical tool known as *systems factorial technology* to answer research questions about the architecture underlying cognitive task. This tool builds on early methodologies used to estimate the time required to complete specific cognitive subtasks. Precisely, it integrates the *subtractive method* (Donders, 1969) and the *additive method* (Sternberg, 1969) to create a double factorial paradigm, from which the data can be analyzed to assess the cognitive architecture of a task. In this paradigm, participants must process two types of information and the researcher manipulates the saliency of these two types of information independently one from another. Using systems factorial technology, it is possible to diagnose whether these two types of information are processed in serial, in parallel or in coactive, and whether the stopping rule used by the participant is self-terminating or exhaustive. A serial architecture means that participants process the two information in succession. A parallel architecture means that participants process the two information simultaneously. The stopping rule indicates when the processing is completed, either as soon as one information is processed (self-terminating) or when both information is processed (exhaustive). In all cases, it posits that the subprocessing unit responsible to process a certain type of information can trigger a response independently of the other subprocessing units. Finally, a coactive architecture means that the subprocessing units associated with each information feed their accumulation to a decision unit, which combines information and trigger a response. Consequently, it does not have a clear stopping rule.

In the second study of this thesis, participants were asked to complete a *Same-Different* task and to base their judgment on two types of information: similarity and clarity. Concretely, they had to categorize the second stimulus 1) as *Same* if it was physically similar or identical to the first stimulus, and if it was slightly blurred or clear, or 2) as *Different* if it was either physically dissimilar or very blurred. Analyzing the data with systems factorial technology allowed me to determine if similarity and clarity were processed in serial, in parallel or in coactive. In other words, it allowed me to test the hypothesis that information related to similarity is processed in preference over other types of information.

The article reporting the results of the second study is currently under review by *Acta Psychologica*. Chapter 4 of this thesis is the submitted version of this article.

Study 3 – Local Processing Architecture and Processing Capacity

A recurring idea for the explanation of the fast-same effect is the possibility that *Same* responses and *Different* responses are the result of distinct forms of treatment. For instance, the dual-process model states that *Same* responses originate from a fast treatment of a chunked version of the stimulus, whereas *Different* responses originate from a slower, but more thorough treatment. Alternatively, the attention-driven single-process model suggests that *Same* responses are the result of a pre-attentive treatment, whereas *Different* responses require attentional resources. The facilitation model mentions that *Same* information is accumulated faster than *Different* information because of stimulus repetition. These models all have alternative perspectives on what distinguishes the processing of *Same* information from the processing of *Different* information, but none truly formalize this distinction.

Concretely, these models propose that elements of information are processed in a unique way when they are *Same* and in a distinct way when they are *Different*. Again, this is a question of cognitive architecture. Imagine a scenario in which the following letters are presented to a participant as a first stimulus: “J C V D”. The participant is then presented with the same letters “J C V D”, processed the second stimulus and respond *Same*. The next trial, the participant is shown “J C G P” followed by “J C V D”. They process the exact same letters as in the previous trial, but this time, their response time is slower. What was different? There are three possibilities. First, nothing was different in the treatment; participants were biased for the *Same* response and that explains the discrepancy. The second possibility is that in the first trial, the letters were processed simultaneously (in parallel, or chunked), whereas in the second trial, the letters were processed one after the other (in serial, or analytically). In the third and final study of this thesis, I focus on this research question: what concretely differs between the processing of *Same* and *Different* information? To answer this research question, I focus on two aspects: the respective cognitive architecture and processing capacity underlying *Same* and *Different* responses.

Again, systems factorial technology is a useful diagnosis tool to answer these research questions. Like in the second study, I used a double-factorial paradigm, but instead of manipulating global dimensions (i.e., similarity and clarity), I manipulated local elements, here, the letters composing the second stimulus. In this study, there were always two letters presented in the first stimulus (a leftmost and a rightmost letter). In the second stimulus, each letter was manipulated in three possible conditions, independently from the other: either it was clear, blurry or absent. Blurring one letter independently from the other letter is necessary to assess the cognitive architecture, whereas removing a letter is necessary to assess the efficiency at which participants processed the information (capacity).

The methodology and plan of analysis of this study were approved by *Attention, Perception, & Psychophysics* as the Stage 1 of a registered report. Therefore, the journal had *in-principle* accepted to publish the results regardless of the outcome. The full report (Stage 2 of the registered report) is now published. Chapter 5 of this thesis is the accepted Stage 2 version of this article.

Chapter 3 – Study 1: The Fast-Same Effect of an Exclusive-OR Task

Abstract

Participants are faster to decide that two stimuli are identical than to decide that they are different. Opposing theories suggested that this fast-same effect is either due a) to a response bias towards similarity or b) to facilitation caused by the repetition of the stimuli attributes. Although both theories predict the fast-same effect in a conventional Same-Different task, they make distinct predictions for tasks in which response bias is removed. In such tasks, the bias theory predicts that the fast-same would disappear whereas the facilitation theory predicts that the fast-same would remain. We tested those hypotheses using a Same-Different task in which participants had to indicate if all the attributes of the stimuli were matching or all were mismatching by pressing one response key, or if some attributes were matching and some were mismatching, by pressing another response key. We call this an Exclusive-Or Same-Different task. Results show that participants were much faster in the *All-Matching* condition compared to the *All-Mismatching* condition, therefore supporting the facilitation theory. A fit of the Linear Ballistic Accumulator model to the observed data provide additional supports that the fast-same effect is not caused by bias, but by a faster accumulation rate of evidence in the *All-Matching* condition.

<u>Reference of the article</u>

Goulet, M.A., & Cousineau, D. (in press). The Fast-Same Effect of an Exclusive-OR Task. *Journal of Experimental Psychology: Human Perception & Performance*.

Introduction

A well-known, often replicated finding in Same-Different tasks is that participants identify *Same* stimuli faster than *Different* stimuli (e.g., Bamber, 1969; Egeth, 1966; Nickerson, 1965; Posner & Mitchell, 1967). Intuitively, one would expect the opposite: whereas identical stimulus pairs require participants to exhaustively process all the attributes of the stimuli, *Different* stimulus pairs only require locating a single mismatching attribute. The logical implication is that participants should not be slower in the *Different* condition, but empirical results show otherwise. Despite over 50 years of research, the many reviews on the fast-same effect have failed to thoroughly explain why it occurs (Farell, 1985; Krueger, 1978; Nickerson, 1973b, 1978; Ratcliff, 1978, 1981, 1985; Sternberg, 1998). Although this effect is also observed when presenting the two stimuli simultaneously (Chignell & Krueger, 1984; Proctor, 1981; Proctor & Healy, 1985; Proctor & Rao, 1983b; Proctor et al., 1984), the focus of this study will be on the version of the task in which the stimuli are presented in succession.

Some authors proposed that a dual-process model can explain the fast-same effect (Bamber, 1969; Bamber et al., 1975; Decker, 1974; Derks, 1972; Nickerson, 1973a; Tversky, 1969) while others argued that a more parsimonious single-process explanation can also predict this result. We divide these single-process explanations in two categories: 1) the fast-same is caused by the participant's bias; and 2) the fast-same is caused by the stimulus' characteristics.

The participant's bias explanation implies that the participants lean towards *Same* responses compared to *Different* responses prior to the presentation of the stimuli (Irwin et al., 2001; Ratcliff, 1978; Ratcliff & Hacker, 1981; Ratcliff et al., 1989; D. A. Taylor, 1977). Indeed, humans might have an inherent preference for similarity (for example, to detect a familiar environment, or familiar faces). Models of information accumulation (e.g., the drift-diffusion model, Ratcliff, 1981) can easily illustrate this bias. In the drift-diffusion model, the decision between *Same* and *Different* is depicted by an accumulator drifting towards one of two thresholds (one for each response). Whenever the accumulator reaches a threshold, the participant provides the corresponding response following a motor decision time. The bias explanation suggests that participants start their accumulation process closer to the *Same* threshold than the *Different* threshold.

Using best-fitting parameters, Ratcliff (1985) showed that the bias parameter (also known as the starting point parameter) could explain the fast-same effect (see Proctor, 1986, for a thorough discussion of the model). However, Proctor (1986), referring to previous studies in which he manipulated the ratio of *Same* vs. *Different* responses (Proctor & Rao, 1983b; Proctor et al., 1984) noted that the fast-same effect was disrupted only in extreme conditions (when 80% of the trials are different).

Proctor proposed an alternative, whereby the fast-same effect is not caused primarily by bias but by the characteristics of the stimuli. He argued that the encoding of the second of two displayed stimuli is facilitated because the exact same attributes that were presented in the first stimulus, displayed earlier, are repeated (Proctor, 1981; Proctor & Rao, 1983a). In other words, participants are more sensitive to *Same* stimulus pairs. The facilitation explanation finds support in experiments which manipulated the degree of similarity between *Same* stimuli. For example, even though the fast-same is still observed when the stimuli are matching nominally but not physically (i.e., a lowercase “a” and an uppercase “A”), the effect is diminished (Eviatar et al., 1994;

Kinoshita & Kaplan, 2008; Posner & Mitchell, 1967; Proctor, 1981; Walker & Cousineau, 2019).

The bias and the facilitation explanations make similar predictions with regards to the fast-same effect, mainly because this is the effect they aim to predict. However, their conceptualization of the origin of the effect differs.

In the current study, we use an experimental design which yields distinct predictions for both explanations. The participants are asked to compare the stimuli like in a standard Same-Different task with sequential presentation of stimuli. However, instead of responding whether the stimuli are *Same* or *Different*, they must indicate if the attributes are all matching using a specific response key, if the attributes are all mismatching using the same response key or if some attributes are matching and some are mismatching using a second response key. We call this task an exclusive-OR Same-Different task. Using the same response key to respond *All-Matching* and *All-Mismatching* equalizes any potential bias for a response over another.

With this design, the bias explanation predicts that participants will take the same amount of time to respond *All-Matching* and *All-Mismatching*. Indeed, if the participants are biased for that response key, they will be equally biased in these two conditions. Conversely, the facilitation explanation predicts that participants will still manifest a fast-same effect, i.e., the response times in the *All-Matching* condition will be faster on average than in the *All-Mismatching* condition.

These hypotheses were preregistered using the *As Predicted* form, available on the OSF project page of this study: https://osf.io/qdk8z. This page also provides access to the raw data and analysis script used to report the main results in this article.

Methods

The experiment has been approved by the University of Ottawa's Research Ethics Board. Participants were compensated $8 CAD for their participation in the study. They received a debriefing at the end of the experiment describing the main objectives and hypotheses of the study.

Participants and Power Analysis

The participants were twenty undergraduate students (age 18-32; 9 females, 11 males) of the University of Ottawa, recruited by the experimenter (this is the preregistered sample size). Although this is, to the best of our knowledge, the first experiment using an exclusive-OR Same-Different task, we estimated the expected fast-same effect size to be d = 0.40, based on previous published and unpublished studies conducted in our laboratories on the (standard) Same-Different task (Bamber, 1969, 1972; Farell, 1988; Harding, 2018; D. A. Taylor, 1976a). Accounting for the fact that the participants' measures are aggregated over a large number of trials (200 trials per condition), twenty participants represent a power of about 0.95 to detect such a fast-same effect (Goulet & Cousineau, 2019b).

Stimuli

The stimuli were arrays of random combination of either two uppercase letters (sampled from B, C, D, F, J, K, L, N, S, T, V, and Z) or two digits (sampled from 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9). We used two different types of stimuli to show that the results are not bounded to a single stimulus type only. From now on, we use *attribute* to refer to a single letter or digit within an array. In a single array, no attribute is repeated. The task of the participant is to compare a pair of arrays and determine if the list of two attributes are all matching or all mismatching (if so, they press one response key), or if one attribute is matching and one is mismatching (if so, they press the other response key). Matching attributes always have the same position in the array.

The stimuli had a luminance of 44.33 cd/m^2 and the background had a luminance of 0.09 cd/m^2.

Procedure

The stimuli were displayed on a CRT monitor (1024×768 pixels; refresh rate of 85 Hz) situated at approximately 50 cm of the participant. A trial started with a fixation point located at the centre of the screen, for a duration of 506 ms, followed by the first array of two attributes, S_1 (horizontal visual angle display of about 5.2°), presented slightly above the fixation point (visual angle of about 4.0°) for 400 ms, after which it disappeared. An inter-stimulus interval of 400 ms, during which a blank screen is presented to the participant, separated the presentation of S_1 from the second array of attributes, S_2, which appeared slightly below the centre of the screen (visual angle of about -4.0°) for 5000 ms, or until the participant provided a response.

The participants were instructed to press either the *Control* or the *Enter* keys (at each extremity of the keyboard) to respond. The assignment of the response key is controlled for participants' handedness. Half of the participants answered *All-Matching* and *All-Mismatching* with their dominant hand, and *Some-Matching-And-Some-Mismatching* with their non-dominant hand, and vice versa for the other half.

Design

There are four conditions in this study: (1) the two attributes are matching (*All-Matching*); (2) the two attributes are mismatching (*All-Mismatching*); (3) the left attribute is matching and the right is mismatching (*Match-Mismatch*); and (4) the left attribute is mismatching and the right is matching (*Mismatch-Match*). Each condition is replicated for 200 trials. All trials from the four conditions are intermixed randomly so that participants cannot predict condition of a trial beforehand.

The task is divided into four blocks. At the end of each block, participants are offered to take a short break. The type of stimulus (letters or digits) is constant in each block and alternate after every break. The type of stimulus of the first block is counterbalanced across subjects.

Results

In this section, we report the outliers analysis first, the result of the confirmatory, preregistered analysis next and finally, additional exploratory analyses.

Data Screening and Outliers Analysis

A first screening of the data shows that one participant performed particularly poorly in two conditions. This participant had an accuracy of 57.5% and 39.0% in the *All-Matching* and *All-Mismatching* conditions respectively. These scores are 2.97 and 2.36 standard deviations below the sample's mean. We decided to remove this participant from further analyses.

For the remaining participants, we have 46 missing responses because the participants did not respond within 5000 ms. We also removed 16 trials for which response times were below 100 ms and 23 trials for which response times were above 4000 ms because it is unlikely that participants performed the task as asked for these specific trials. In total, 85 trials were removed out of 15,200 (0.56% of the data).

We plotted the main results in Figure 3.1.[7] As seen, the stimulus type has no effect on the mean response time (RT) in any condition. We therefore combine the stimulus type in our subsequent analyses.

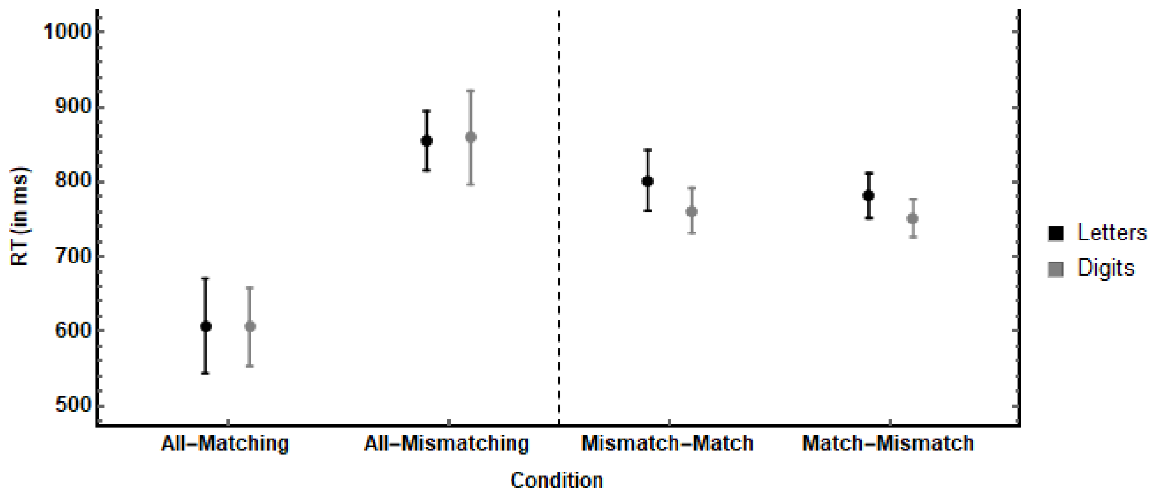


Figure 3.1. Mean RTs in all experimental conditions. The black and gray dots respectively correspond to trials in which letters and digits were used as attributes. The dashed line represents the response key separation; to its left, the All conditions, and to its right, the Some conditions. Error bars are difference- and correlation-adjusted 95% confidence intervals.

Confirmatory Analysis

We tested the main hypothesis of this study using a paired sample *t*-test between the *All-Matching* and the *All-Mismatching* conditions. In these conditions, the participants pressed the exact same response key to provide their response. As can be

---

[7] All plots include difference- and correlation-adjusted 95% confidence intervals. These adjustments reflect those used in the reported statistical tests (difference-adjustment for comparing multiple groups and correlation-adjustment for comparing repeated measures; Baguley, 2012; Cousineau, 2017).

expected from Figure 3.1, there is a very large difference between the two conditions (t(18) = 7.91, p < .001, d = 1.8, BF_{10} = 86,878.31).[8] Indeed, the mean RT in the *All-Matching* condition is 250 ± 32 ms faster than the mean RT in the *All-Mismatching* condition.[9]

The results of this study strongly support the facilitation explanation. The discrepancy between the *All-Matching* and *All-Mismatching* conditions cannot be explained by a response bias as participants used the same response key to provide their response. Conversely, the facilitation explanation predicted – successfully – that the comparison process would be facilitated in the *All-Matching* condition. This suggests that the fast-same effect is the result of a repetition of the stimulus characteristics, rather than an inherent bias for the *Same* response.

Exploratory analyses

An unexpected result is that the fast-same effect is larger than anticipated. Typically, the effect in standard Same-Different task is about 40 ms with a standard deviation of the mean difference of about 100 ms. In the present study, the fast-same effect is over six times larger (250 ms) and the standard deviation has almost doubled in size (191 ms). Undoubtedly, participants performed the Same-Different task differently in this exclusive-OR version than under its more standard version. The fact that the effect

---

[8] The Bayes Factor corresponds to the amount of evidence in favour of the alternative hypothesis (the mean RTs in the two conditions are different) relative to the evidence in favor of the null hypothesis. It has been calculated using a Cauchy distribution as a prior (location: 0.4, scale: 0.707).

[9] All means and means differences are presented with plus or minus one standard error of the mean or plus or minus one standard error of the mean difference.

size is larger than expected does not change the conclusion of the confirmatory analyses, quite the contrary because the bias explanation predicts a null effect. That said, we wish to investigate possible reasons for the large effect size using additional analyses.

We perform three more analyses. First, we investigated the RTs in the *Some* conditions. Second, we examined the accuracy of participants in all the conditions. Third and last, we used a theory-driven model, the linear ballistic accumulator (LBA; Brown & Heathcote, 2008), to estimate parameters showing how much participants were biased, their information accumulation rate and their non-decision time in each condition.

***Some* conditions RTs**

We see in Figure 3.1 the mean RTs of the *Some* conditions. Whereas in the *All* conditions, we observe a large discrepancy between the mean RTs, in the *Some* conditions, the difference is only 14 ± 7.7 ms, $t(18) = 1.9$, $p = .077$, $d = 0.4$, $BF_{10} = 1.29$.

If we combine the two *Some* conditions together, participants averaged a response latency of 774 ± 47 ms, compared to 607 ± 29 ms in the *All-Matching* condition and 857 ± 55 ms in the *All-Mismatching* condition. In other words, the mean RTs difference between the *Some* and the *All-Matching* conditions is 167 ± 23 ms, which is twice as large as the mean RTs difference between the *Some* and the *All-Mismatching* conditions (83 ± 16 ms). This could be explained by processing facilitation alone. Perhaps the facilitation is greater in this version of the task compared to the standard variation of the *Same-Different* task. However, we consider another possible alternative to address this discrepancy in effect size.

The pattern of mean RTs described above suggests a sequential examination of the outcome alternatives (Cousineau, 2004). Even though there are two possible responses (*All* and *Some*), participants might approach the task with three possible outcomes in mind: are the attributes 1) all matching? 2) partially matching? / partially mismatching? or 3) all mismatching? If such is the case, the rundown of a trial would be as follows. First, the participant seeks to determine whether the two attributes are matching. If this is the case, the participant responds *All*. If this step is inconclusive, the participant enters the second step, determining if one attribute is matching and one is mismatching. This step adds about 160 ms to response times. When one is identified as a match and the other as a mismatch, the participant responds *Some*. If this is not the case, the participant finally enters the third phase, in which upon a supplementary 80 ms, the participant responds *All*.

Accuracy

If participants indeed ponder response alternatives sequentially in the current task, it should also be reflected in their accuracy scores. Say that on a given trial, the stimuli are identical (*All-Matching* condition). If participants processed the stimulus and found after the first step that the two elements composing the stimulus are matches, they can response accurately *All*. However, if the participants did not come to that conclusion, they still have not committed to responding *Some*; they simply proceeded to the next step (is there some matches and some mismatches?). This gives enough time to the participants to realize their mistake and respond *All*. The same situation is valid for the *Some* trials. If the participants failed to accumulate enough evidence to respond *Some* in the second step, they can always rectify their response during the third step. However, in the case of an *All-Mismatching* trial, the correct response only occurs at the third and final step, meaning that the participants have no opportunity to rectify their response afterwards.

Therefore, the sequential examination of the outcome alternatives predicts that the accuracy should be lower in the *All-Mismatching* condition compared to the other condition.

We tested that prediction using the same format as in Figure 3.1, that is, we plotted the accuracy scores of the participants in every condition in Figure 3.2. Again, we see only small differences between the stimulus types. We henceforth combine these scores to ease the interpretation of the results. As seen, the *All-Mismatching* condition has the lowest mean accuracy scores with 81% ± 3%. The other conditions have similar accuracy scores, their means lying between 89% and 91%. When compared to the *All-Mismatching* condition, the differences expressed as Cohen's *d* are approximately equal to 1.0, $p < .005$ and $BF_{10} > 100$ using a null prior. This further supports the idea that participants ponder the response alternatives serially.

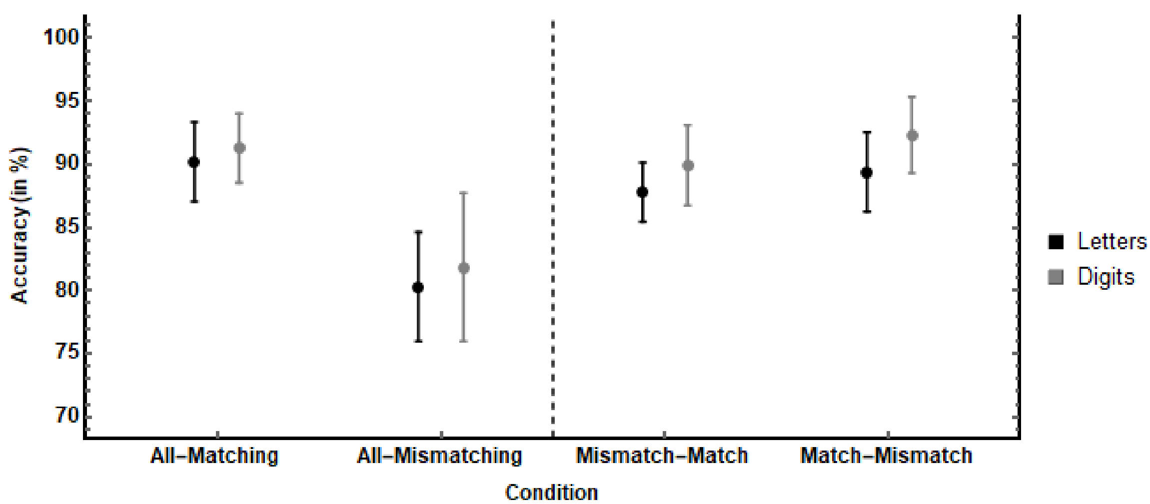


Figure 3.2. Mean accuracy scores in all experimental conditions. The black and gray dots respectively correspond to trials in which letters and digits were used as attributes. The dashed line represents the response key separation; to its left, the All conditions, and to its right, the Some conditions. Error bars are difference- and correlation-adjusted 95% confidence intervals.

Model Fitting

The final exploratory analysis focuses on the estimation of the bias and the accumulation rate parameters obtained through the LBA. This theoretical model posits that participants likely accumulate information about the stimulus (or evidence in favour of a response) over time at a certain rate. This evidence accumulation framework estimates the values of six parameters, which are related to specific cognitive constructs. It posits that participants must accumulate a certain amount of evidence to respond, which is represented by a decision threshold. This could represent, for example, the brain activity required to activate a motor response. Second, the LBA states that participants gather those evidences by processing the stimuli and accumulating evidence in favour of the two response alternatives. In our case, participants would accumulate evidence for the *All* and for the *Some* response, each *accumulator* gathering evidence at a certain rate.

In each condition, the LBA represents those accumulators as the correct response accumulator and the incorrect response accumulator. The rates of accumulation for these two accumulators vary between trials, which is depicted by the accumulation rate variability parameter. The average accumulation rates of the correct response and incorrect response accumulators are depicted by two parameters. The distribution of accumulation rates over the course of the entire experiment follows a normal distribution. Additionally, at the beginning of the trial, prior to processing the stimulus, participants might have pre-emptively accumulated evidence in favour of a response, without having seen S_2, resulting in a bias. The maximum possible response bias on a given trial is depicted by the starting point parameter of the LBA. The distribution of starting points over the course of the entire experiment follows a uniform distribution. Finally, the LBA accounts for the time taken by participants to generate a motor response, after the decision process is done, which is represented by the non-decision base time parameter.

The advantage of using this model is that we can estimate the value of the starting point and the correct response accumulation rate in each condition and for each participant (for model fits of the classical version of the *Same-Different* task, see Goulet & Cousineau, 2019a and T.-Groulx et al., in 2019). We can relate those parameters to evaluate how much response bias and accumulation rate explain the RT difference between the condition, specifically the fast-same effect.

We used a maximum likelihood estimation procedure using the simplex method to fit the model to all trials (Nelder & Mead, 1965).[10] We estimated the best-fitting parameters for each participant independently. We also estimated the best-fitting parameters for four experimental conditions: *All-Matching*, *All-Mismatching*, *Match-Mismatch* and *Mismatch-Match* (we merged the letters and the digits trials). For each condition, we estimated the value of five parameters: 1) starting point, 2) accumulation rate of the correct response accumulator, 3) accumulation rate of the incorrect response accumulator, 4) inter-trial variability of the accumulation rate, and 5) non-decision base time. A sixth parameter, the decision threshold, was fixed at 800 evidences for all the

[10] We also fitted the model using a different algorithm, the differential evolution method (Storn & Price, 1997). This method takes a considerably longer time to fit the results (about 3 hours per participant) and failed to find a solution for three participants (even after changing the starting values of the algorithm). The fit for the other sixteen participants was slightly better (average log-likelihood per datum of -6.67). That said, we replicated Figure 3.3 and 3.4 with the new parameters and found virtually no difference. This means that the parameter values reported in this article are robust to the estimation algorithm.

conditions. Because the conditions are randomly presented, participants cannot set their decision thresholds in advance. For this reason, we decided to use a fixed value for the decision threshold, which also acts as a scaling parameter.[11] Therefore, we fitted four experimental conditions for each participant using 20 free parameters and 1 fixed parameter.

Three constraints were imposed on the parameter values during the search. The first constraint was that the bias parameter had to be smaller than the threshold. The second constraint was that the non-decision time had to be greater than zero, but smaller than the fastest observed RT. The accumulation rates were constrained to be any positive value.

With the LBA, we were able to accurately replicate the trends observed in the data (mean log-likelihood per datum = -6.76), as can be seen in Figure 3.3, albeit it predicts faster mean RT in the *Some* conditions. This is probably because the model tries to account for the rare very fast trials (< 300 ms) in these conditions. Nevertheless, the model predicts the fast-same effect that is observed in the data.

The best-fitting parameters of the LBA are plotted in Figure 3.4. First, we test the effect of the condition on the starting point parameter using a repeated-measures one-way ANOVA. The results indicate that there is a large effect of the conditions on the starting point, $F(3, 54) = 14.68$, $p < .001$, $\eta_p^2 = .449$, $BF_{10} = 44,465$. Post-hoc tests show that the

[11] Although we only need to fix one parameter in one condition to scale all the free parameter values (Donkin et al., 2009), we decided to fix threshold parameter to the same value in every condition. We took this decision because we did not want to add more free parameters to the model and because we did not have any theoretical justification to let those parameters be free.

effect is explained by higher starting points when participants responded *All* compared to *Some*. However, within the response type, the starting points are similar. Indeed, the parameter values are about the same for the *All-Matching* and the *All-Mismatching* conditions (mean difference of 7.79 ± 54.77 evidence, $p_{\text{Bonferonni}} > .999$, $BF_{10} = 0.24$), and for the two *Some* conditions (mean difference of 70.36 ± 40.79 evidence, $p_{\text{Bonferonni}} = .610$, $BF_{10} = 0.82$).

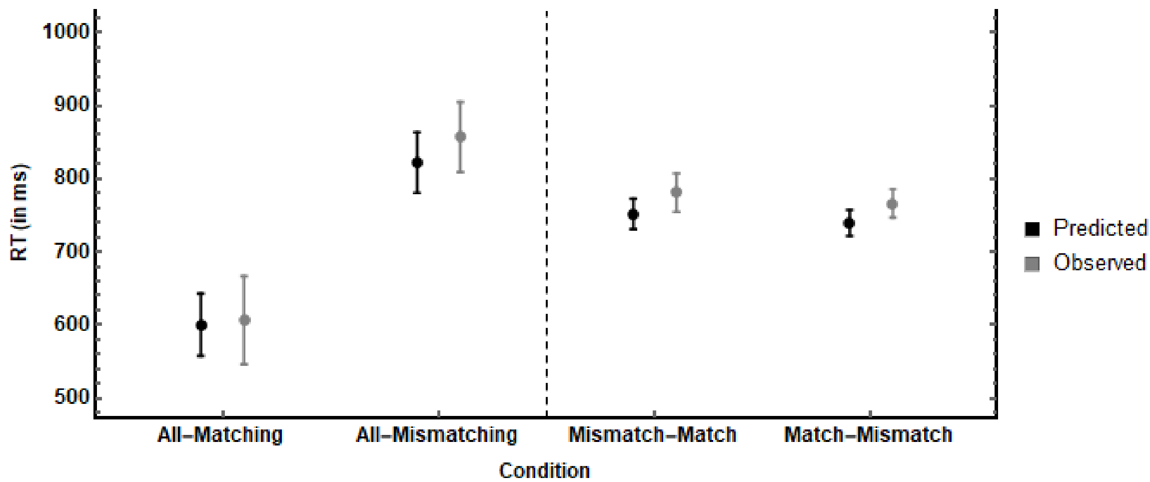


Figure 3.3. Comparison of the LBA predicted RTs (black dots) and the observed RTs (gray dots) for four experimental conditions. To ease readability, we merged trials that used letters and digits. The dashed line represents the response key separation; to its left, the All conditions, and to its right, the Some conditions. Error bars are difference- and correlation-adjusted 95% confidence intervals.

These results suggest that participants were more biased to respond *All* than *Some*. They also suggest that in the *All* condition, participants did not benefit from a response bias advantage in the *All-Matching* compared to the *All-Mismatching* condition. This shows that 1) our manipulation was effective and 2) that there is a response bias, but it is not the cause of the fast-same effect because the same response bias is found for the *All-*

mismatching condition as well. Consequently, the fast-same effect is not caused by response bias.

Second, we conducted the same analysis but on the accumulation rate parameter. Again, we observe a large effect of the conditions, $F(3, 54) = 14.12$, $p < .001$, $\eta_p^2 = .440$, $BF_{10} = 24{,}392$. Post hoc tests indicate that participants accumulated information at a faster rate in the *All-Matching* condition compared to the other conditions, mainly the *All-Mismatching* condition (mean difference of 0.43 ± 0.06 evidence per ms, $p_{\text{Bonferroni}} < .001$, $BF_{10} = 13{,}984.1$). These results imply that the LBA attributes the fast-same effect to a faster accumulation rate (and not to larger response bias) in the *All-Matching* condition compared to all the other conditions, including the *All-Mismatching* one. This is congruent with our interpretation of the confirmatory analysis and further supports the facilitation explanation.

Finally, we performed the same analysis on the non-decision base times. We observed a medium effect of conditions on the base time, $F(3, 54) = 6.47$, $p < .001$, $\eta_p^2 = .265$, $BF_{10} = 44.1$. Post-hoc tests show that the effect is explained by the *Match-Mismatch* condition having a slightly faster base time than the *All-Mismatching* (mean difference of 139.4 ± 38.0 ms, $p_{\text{Bonferroni}} = .011$, $BF_{10} = 22.78$) and the *All-Matching* (mean difference of 91.1 ± 28.4 ms, $p_{\text{Bonferroni}} = .024$, $BF_{10} = 9.48$) conditions. We are unsure how to interpret those results, as they do not reflect anything that was observed in the RT data. For this reason, and because the base times estimations are highly variable, we do not further linger on this effect.

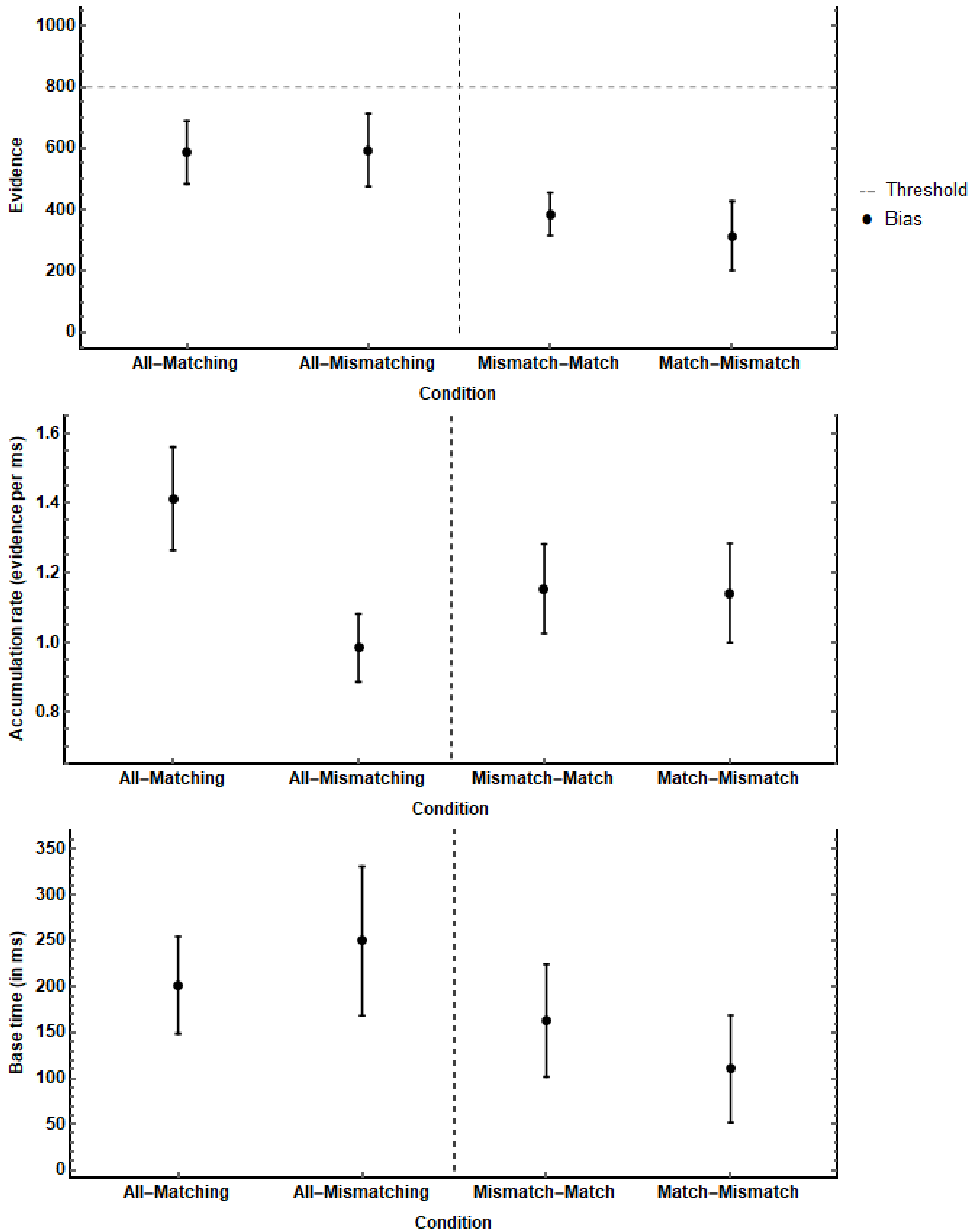


Figure 3.4. Mean estimated parameter values (top – bias and threshold; middle – accumulation rate for the correct response; bottom – base times) for four experimental conditions. The letter and the digit conditions have been aggregated. The vertical dashed line represents the response key separation; to its left, the All

conditions, and to its right, the Some conditions. Error bars are difference- and correlation-adjusted 95% confidence intervals.

We also fitted another sequential sampling model, the Ratcliff's Drift-Diffusion Model (DDM; Ratcliff, 1981) to the observed data. We used a similar technique as with the LBA (maximum likelihood estimation using the simplex method; all parameters were free). However, we found that the fit was inadequate and judged that it qualitatively did not replicate the observed results (mean log-likelihood per data = -7.09). Among others, it failed to predict the mean RT discrepancy between the *All-Matching* and the *All-Mismatching* conditions, the cornerstone result of this experiment. For this reason, we did not pursue further investigation using the DDM. Nonetheless, for sake of transparency, we mention this failed attempt (Simmons et al., 2011).

Discussion

The main finding of this study is that the fast-same effect is likely caused by a processing facilitation rather than a response bias. These two competing explanations made distinct predictions about the mean RTs discrepancy between the *All-Matching* and *All-Mismatching* conditions in an exclusive-OR task. The results showed unambiguously that the *All-Matching* condition still yielded faster mean RTs than the *All-Mismatching* condition even though both conditions used the same response key. Therefore, the results support the facilitation explanation.

The facilitation explanation was also supported in our exploratory analysis using the LBA. Whereas *t*-tests and analyses of variance compared mean RTs in each experimental condition, maximum likelihood parameterization uses the entirety of the trials to fit the data. More importantly, these parameters are tied to theoretical concepts. In the present experiment, we estimated four parameters, including two that are central to our research question: bias and accumulation rate. Whereas the bias explanation would predict higher bias in the *All-Matching* condition, the facilitation explanation would predict a higher accumulation rate in this condition. As described in the *Results* section, the latter was observed, further supporting the facilitation explanation.

One could argue that even though the fast-same effect seems to be caused by facilitation, it still does not provide a strong argument against the response bias. Perhaps response bias only contributes to the fast-same effect in the classical format of the *Same-Different* task and not in the one used in the current study. Whereas this possibility is real, we argue that it is unlikely. If the fast-same effect is the result of both response bias and facilitation in the classical *Same-Different* task, the fast-same effect should be larger than the one observed in the current study where only facilitation seemed to be in operation. Yet, the fast-same effect reported in this study is about six times *larger* than the one typically observed in the *Same-Different* task. When interpreting the results using effect sizes, it is difficult to explain how response bias would contribute to the fast-same effect.

Instead, we attribute the large effect reported in this study to participants pondering response alternative serially. We must keep in mind that this might have affected the estimation of parameters in the LBA fit. For instance, the discrepancy between accumulation rates in the *Same* condition vs. the other conditions might be inflated due to the large fast-same effect observed in the data. It remains unclear whether response facilitation is increased in the exclusive-OR task compared to the more traditional *Same-Different* task. For instance, Goulet and Cousineau (2019a) found a much smaller effect of the accumulation rate when fitting the LBA to traditional *Same-Different* data. Future research will be able to address this question.

The locus of the facilitation effect is related to the repetition of the attributes. If participants know what attributes to expect (as provided by S_1), they may create a mental representation of these attributes, which acts as a *boost* for processing matching information (Farell, 1984, 1988; Kinoshita & Kaplan, 2008). Some authors suggested to describe this as a form of priming induced by the mental representation in working memory (Jacob et al., 2013; Kinoshita & Kaplan, 2008). To avoid confusion with priming caused by perceptual memory (such as the iconic memory), we prefer to use processing facilitation to describe this theoretical approach, as proposed by Proctor (1981, 1986).

The idea that mental representations are critical to the fast-same effect is not novel. Farell (1985) hypothesized that when participants create mental representations of the stimuli, they can more easily map matching attributes one-to-one. The representation of mismatching attributes is different, because participants cannot anticipate what a mismatching attribute looks like (i.e., in this task, it is not a precise letter or digit, but any letter or digit left in the pool of unused symbols). Therefore, the mapping of mismatching attributes is many-to-one. This view is also supported by studies that asked participants to ignore the relative position of the attributes in the string (Proctor & Healy, 1985, 1987; Sinha & Glass, 2017). In these studies, participants' performances in *Same* trials decreased (slower RTs and higher error rates) as the positional displacement increased, but the fast-same effect was nevertheless observed. This suggested that participants might have some flexibility with their mapping, as long as they know what the attributes composing a *Same* stimulus look like. According to these authors, when the position of the attributes is irrelevant, participants still benefit from encoding facilitation, but it is not as large as when the position of the attributes is kept constant.

Farell, on the contrary, did not attribute the fast-same effect to faster accumulation rates, but to a switch in processing phase. Because one-to-one mapping is easier, participants likely perform such comparisons first, then switch to many-to-one mapping, if necessary. As he mentioned, this conceptualization is akin to the dual-process approach put forth by Bamber (1969).

A characteristic of this dual-processing mode theory is that it postulates serial processing. Our analysis of the *Some* conditions and of the effect sizes between all the conditions suggested that participants manifested such serial treatment in our study. We also found support for the serial treatment with the accuracy scores. Participants made more errors in the *All-Mismatching* condition, the last step of the serial treatment. Because this is the last step, participants do not have an extra branch to rectify their mistake in case of an erroneous conclusion. Errors in the first two phases are caused by a lack of signal; the threshold was not reached. However, errors in the *All-Mismatching* condition means that the signal in the second branch was too strong. According to this model, participants would have to wait (about 80 ms) to provide a *All* response, enough time for the *Some* threshold to be – erroneously – reached.

Overall, the results of this experiment can be explained by a sequential examination of outcomes: participants evaluate each response alternatives in succession. In the case of the exclusive-OR Same-Different task, participant must evaluate three response alternatives instead of two, even though there are only two possible responses.

That said, a main limitation of this conclusion is that it was made using exploratory analyses. Indeed, the current experiment was not specifically designed to directly assess the seriality of the treatment (or the *architecture* of the underlying cognitive system). Better experimental designs can diagnose cognitive architectures using logical-rule models (Bushmakin et al., 2017; Eidels et al., 2010) or a double factorial paradigm (Goulet & Cousineau, 2020). These approaches can be analyzed with a statistical tool known as Systems Factorial Technology (Daniel R. Little, Altieri, et al., 2017; Townsend & Nozawa, 1995).

Another critical element of this sequential examination of outcomes approach is that it posits that participants only gather matching information, not mismatching information. From a facilitation perspective, it means that accumulating information is faster than waiting for a threshold not to be met. It is logical: if participants want to be accurate, their "*did not cross the threshold*" timer must be long enough so that the *matching information* accumulator has a chance to reach the threshold. However, one could argue that participants also gather mismatching information (Hyun et al., 2009). Such information could serve as an inhibition of the *Same* response (Proctor, 1981).

In summary, the results presented here support the theory that the fast-same effect is caused by a facilitated treatment of matching information. They also reject the theory that the fast-same effect is caused by an inherent bias for the *Same* response. Indeed, the fast-same effect was still observed even though we removed any form of response bias (this was also supported by the LBA fit). Additional exploratory analyses showed that participants likely used a sequential examination of outcomes strategy to complete the task. In other words, the participants pondered the response alternatives serially until enough information was accumulated to respond. However, it remains unclear if the participants only accumulated information about matching attributes, or if they also concurrently accumulated information about mismatching attributes.

Chapter 4 – Study 2: The Cognitive Architecture of Processes Responsible to Assess Similarity and Clarity in a Comparison Task

Abstract

When asked to compare two stimuli, participants are on average faster to respond *Same* than *Different*, an effect coined the *fast-same*. The dual-process theory argues that information about similarity is processed in priority over any other type of information, causing the fast-same effect. We tested this serial architecture of cognitive processes using a double factorial paradigm, suitable for a Systems Factorial Technology (SFT) analysis. Twenty participants completed a task in which they compared two letters, which were varied on two dimensions: the similarity and the clarity of the letters. Their task was to indicate if the second letter was the *Same* as the second letter (ranging from identical and clear to similar and slightly blurry) or if it was *Different* (if the stimuli were either dissimilar or very blurry). The SFT results show that most participants processed the information in a serial fashion, but in a mixed order. In other words, for some trials, participants processed similarity first, and for some other trials, they processed clarity first. This implies that participants indeed process information serially in the comparison task, but that it does not cause the fast-same effect.

Introduction

Humans compare stimuli presented in close succession with ease and accuracy indicating whether the two objects are the same or are different. Researchers studied this ability extensively since the 1960s, using what they coined a comparison task (or a *Same-Different* task; see seminal papers from Bamber, 1969; Egeth, 1966; Nickerson, 1965, 1967 and Posner & Mitchell, 1967). From the perspective of participants, the task is simple: a pair of stimuli is presented, and they respond whether the two stimuli are identical (*Same*) or if one or many of their attributes differ (*Different*). From the perspective of the researchers, this task can be manipulated on multiple levels.

One manipulation concerns the type of stimulus. Whereas some researchers used geometrical shapes and colours as stimuli (Bindra et al., 1968; Egeth, 1966; Grill, 1971; Hawkins, 1969; Jacob et al., 2013; Nickerson, 1967b, 1967a; R. L. Taylor, 1969), most used letter strings (Bamber, 1969, 1972; Bamber et al., 1975; Bamber & Paine, 1973; Decker, 1974; Farell, 1988; Goulet & Cousineau, 2019a; Proctor et al., 1984; Ratcliff & Hacker, 1981; D. A. Taylor, 1976a; Walker & Cousineau, 2019). There are two main rationales for using letter stimuli. First, participants can easily distinguish letters apart. Second, letters allow researchers to easily manipulate other aspects of the stimuli.

For example, researchers can vary the length of the strings (the number of letters). In that regards, participants take more time to respond when the stimuli contain more letters (Bamber, 1969; Bamber et al., 1975; Farell, 1977, 1988; Hyun et al., 2009; Nickerson & Pew, 1973; D. A. Taylor, 1976a). This is logical: more letters mean more information to process. Using multiple letters as attributes, researchers also showed that participants take more time to respond when there are fewer mismatching letters between the two stimuli. Again, this is logical: if participants only need one mismatching letter to respond *Different*, reducing the number of mismatching letters renders the decision more difficult. This is observed except for one condition: trials with zero mismatching letters (*Same* trials). Whereas one would expect *Same* trials to follow the trend and be the slowest condition, as it requires a fully exhaustive comparison process, participants actually have the fastest mean response times in such trials (Bamber, 1969; Egeth, 1966; Nickerson, 1965; Posner & Mitchell, 1967). This phenomenon is referred to as the *fast-same effect*.

Nickerson (1965, 1967b, 1967a, 1968, 1969) argued that this effect defies the then dominant view that stimuli were processed serially (Atkinson et al., 1969; Sternberg, 1966). Indeed, if it was the case, a serial examination of the stimulus's attributes (i.e., individual letter) would mean that *Same* responses would be slowest, because participants would need to process all the letters before responding.

A first theoretical approach put forth to explain the fast-same effect postulated that *Same* responses could arise from a distinct processor only capable of identity detection (an *identity reporter*; Bamber, 1969). Using associations between pictograms and nouns, Tversky (1969) further emphasized this duality in processing strategies. He argued that participants used a *compare-and-check* process to detect similarity (a holistic comparison), followed by an analytical process (a serial comparison) to detect differences. Beller (1970) proposed that the first processor is pre-attentive, suggesting that participants can respond *Same* without resorting to attention (an idea that will be extended by Farell, 1977, 1984, 1985, 1988).

Following these studies, the dual-process model gained in popularity (Bamber et al., 1975; Decker, 1974; Derks, 1972; Krueger, 1973; Nickerson & Pew, 1973; Silverman & Goldberg, 1975; D. A. Taylor, 1976b). This easy-to-understand solution for the fast-same phenomenon relies on the fact that *Same* responses are on average faster than *Different* responses because participants can respond *Same* using a fast, holistic identity processor, while *Different* responses are the result of a slower, analytical processor. The two treatments are done successively: first, participants process information about global similarity; next, they process local information about differences. That said, this explanation has never been directly tested, due to the lack of analysis tools capable of diagnosing cognitive architecture. This has changed since the introduction of the *systems factorial technology* (SFT; Little, Altieri, Fifić, & Yang, 2017; Townsend & Nozawa, 1995).

The dual-process model was criticized in two ways. First, it struggles to account for the modulation of the fast-same effect by the level of comparison. If participants are asked to compare the stimuli at a nominal level (i.e., associate uppercase and lowercase letters, instead of physically identical letters), the fast-same effect is still observed, but the size of the effect is smaller (Bamber & Paine, 1973; Ben-David & Algom, 2009; Pachella & Miller, 1976; Posner & Mitchell, 1967; Well & Green, 1972) and does not change with extended training (Walker & Cousineau, 2019). It is unclear how the dual-process model can account for this modulation of the fast-same effect.

Second, the dual-process model is not parsimonious: it may overfit the results (two processors to fit two conditions). For this reason, researchers in the late 1970s and early 1980s focused on single-process models of the comparison task. Krueger (1978) suggested that the fast-same effect might be due to a discrepancy in decision thresholds. Because the stimuli are encoded with noise, participants need to encode information multiple times to avoid responding erroneously. However, participants can identify that two letters are matching using fewer iterations, because it is unlikely that noise changes perceived letters into matching letters by chance.

Another single-process explanation for the fast-same effect suggests that *Same* stimuli are encoded faster due to stimulus repetition (Chignell & Krueger, 1984; Proctor, 1981; Proctor & Rao, 1983a, 1983b; D. A. Taylor, 1977). The encoding of repeated information is facilitated (or *primed*), benefiting *Same* responses only. This could also explain why the fast-same effect is modulated by the level of comparison: In *Same* trials, participants benefit from multiple sources of matches (i.e., physical, nominal, phonological, categorical), each contributing to increased encoding speed (Harding, 2018; Kinoshita & Kaplan, 2008; Lupker et al., 2015; Proctor, 1986; Walker & Cousineau, 2019).

A final single-process alternative proposes that participants manifest a bias for positive responses, such as *Same* (Irwin et al., 2001; Ratcliff, 1981; Ratcliff & Hacker, 1981, 1982; Ratcliff et al., 1989). These authors showed that, by varying the bias parameter of the Drift-Diffusion Model (Ratcliff, 1978), the model can replicate the fast-same effect. However, this approach was critiqued due to a lack of empirical and physiological support (Proctor, 1986). As Proctor (1981) argued, the fast-same effect might be partially caused by bias, but bias alone cannot completely explain the effect (Krueger, 1985).

These proposals left the literature on the comparison task with a plethora of models and a lack of consensus (see reviews of the fast-same effect in Krueger, 1978; Nickerson, 1973, 1978; Ratcliff, 1985; Sternberg, 1998). Recently, we tried to test these different fast-same explanations by fitting theoretical models, such as the Drift-Diffusion Model and the Linear Ballistic Accumulation (Brown & Heathcote, 2008), with little success (Goulet & Cousineau, 2019a). Even using simpler single-process models like the EZ Diffusion model (Wagenmakers et al., 2007), the fits were unconvincing (T.-Groulx, Harding, & Cousineau, 2019). However, the model fits of Goulet and Cousineau (2019a) suggested that non-decisional times might explain some of the fast-same effect. Non-decisional times are expected to stay constant across conditions, but could vary if say, distinct processes underlie these conditions (Spieser et al., 2017). Additionally, a recent fMRI study suggested that *Same* and *Different* responses activate separate brain areas (Sinha & Glass, 2017).

In the end, perhaps researchers were too quick at discrediting the dual-process model. These recent studies suggest that it might still be a legitimate model of the comparison task and a good explanation for the fast-same effect. As mentioned earlier, the seriality of the dual-process model has never been directly tested. Therefore, the purpose of this study is to test the cognitive architecture underlying the comparison task. Specifically, if the dual-process model is exact, participants should process information about the similarity of the stimuli in preference over other types of information, such as clarity level. To test this hypothesis, we used a *double factorial paradigm*, suitable for a systems factorial technology analysis (Griffiths et al., 2017; Townsend & Nozawa, 1995; Townsend & Wenger, 2004).

Diagnosing Cognitive Architecture Using Systems Factorial Technology

The double factorial paradigm consists in presenting the participants with two types of information (factors). In our study, the participants are presented with a first stimulus (S_1), which needs to be preserved in working memory, followed by a second stimulus (S_2), which is manipulated based on two factors: similarity (how physically similar the S_2 is to S_1) and clarity (how physically blurry S_2 is). The stimuli are single letters. There are three possible levels for each factor. For similarity, the letter presented at S_2 can either be i) identical to S_1, ii) similar to S_1 or iii) dissimilar to S_1. For clarity, the letter presented at S_2 can either be i) clear, ii) slightly blurred or iii) very blurred. The task of the participant was to indicate whether S_2 was 1) identical or similar to S_1, and 2) clear or slightly blurred (S_1 is always presented clearly). If such was the case, participants would respond *Same*, but if S_2 was either dissimilar to S_1 or very blurry, participants would respond *Different*.

This paradigm is used to determine how participants treated the two factors, that is, what cognitive architecture underlay their processing. To do so, we compute the interaction resulting from the two factors in the *Same* condition. SFT uses both the *mean interaction contrast* (MIC) and the *survival interaction contrast* (SIC) to diagnose three possible cognitive architectures. In our study, the three possible outcomes are that participants 1) process similarity and clarity successively (serial architecture), 2) process similarity and clarity simultaneously (parallel architecture), or 3) pool the information about similarity and clarity into a decision unit responsible to respond (coactive architecture). SFT can also assess the stopping rule of the processing, that is whether participants need to process both factors exhaustively to respond, or if only one suffices.

However, in our study, participants need to identify that both factors are targets to respond *Same*, therefore we assume that participants use an exhaustive stopping rule.

Another way of representing the double factorial paradigm is to see the manipulations as a way to hinder information processing for a certain factor (or *stretch* the decision process; Fifić & Little, 2017) without interference on the processing of the other factor. This is called selective influence and it is a core assumption of SFT. In our study, participants will likely process information about similarity faster when S_2 is identical to S_1 compared to when it is just similar. Also, participants will likely process information about clarity faster when S_2 is clear compared to when it is slightly blurred. We refer to the condition in which both treatments are optimal as the *High-High* condition (HH) and the condition in which both treatments are hindered as the *Low-Low* condition (LL). When one factor is optimal and the other is hindered, we refer to those conditions as *High-Low* (HL) and *Low-High* (LH).

The MIC consists in comparing the mean response times (RTs) of the HH and LL conditions to the HL and the LH conditions. If the treatment is serial, the slowdown caused by the *low* levels should be additive, and the MIC should equal zero. However, if the treatment is parallel or coactive, the slowdown caused by the *low* levels is not additive, and the MIC should differ from zero.

The SIC is a similar measure, but instead of comparing mean RTs, it compares the survival distribution of all the conditions for each unit of time (*t*; for each ms). The SIC(t) curves are also used to diagnose cognitive architecture, as each of the three possible outcomes is associated with a signature SIC(t) curve (Harding et al., 2016; Daniel R. Little, Altieri, et al., 2017; Townsend & Nozawa, 1995). Although the SIC(t) curve is often analyzed qualitatively, we can also use statistical tests to assess cognitive architecture as described in Houpt & Burns (2017) and Houpt & Townsend (2010).

The *Different* conditions can also be used to diagnose information about the order of treatment if the treatment is serial (Fifić, Little, & Nosofsky, 2010; Little, Nosofsky, & Denton, 2011). These conditions are also referred to as *Interior* (I) and *Exterior* (E). There are four possible conditions. When S_2 is dissimilar to S_1, the clarity level can either be clear (Exterior Clarity; E_C) or slightly blurred (Interior Clarity; I_C). Alternatively, when S_2 is very blurry compared to S_1, the similarity level can either be identical to S_1 (Exterior Similarity; E_S) or similar to S_1 (Interior Similarity; I_S).

To see the usefulness of these conditions, assume the dual-process model whereby the participants process information about similarity and clarity in succession (a serial treatment). Using only MIC and SIC, it remains uncertain whether participants always process similarity or clarity first (a fixed order) or if the order of treatment is random. We can answer this question using the Interior and Exterior conditions. Indeed, if participants process information about similarity first, then the clarity level does not matter in trials where S_2 is dissimilar to S_1, because the participants do not need to process the clarity level (a dissimilar letter is sufficient information to respond *Different*). In other words, the response times between E_C and I_C should be comparable. However, if S_2 is identical or similar to S_1, then participants would need to process clarity and determine that S_2 is blurry. Because the time required to complete the treatment of similarity is dependent on the level of similarity (faster treatment when the letters are identical compared to when the letters are similar), participants should have faster response time in E_S compared to I_S.

Methods

The main objective of this study is to determine if participants process information about similarity in priority over other information, such as the clarity of the stimulus. We address this research question in two ways: 1) what is the cognitive architecture underlying a comparison task? and 2) if that architecture is serial, is the order of treatment fixed or random? We use the double factorial paradigm to answer these two research questions and analyze the results using the SFT analysis suite. Analyses are done at a group level and at an individual level.

Stimuli

The stimuli (S_1 and S_2) are sampled randomly from a set of predetermined combinations. Three sets of letter pairs are generated based on their similarity level (Boles & Clifford, 1989): one set contains pairs composed of identical letters, one set contains pairs composed of similar letters (similarity level > 320), and the last contains pairs composed of dissimilar letters (similarity level < 175). These similarity levels were chosen to maximize three parameters: 1) each set should be of about equal size (how many combinations are possible), 2) the diversity of the sets (the number of letters) should be as large as possible and 3) the distance between the similarity levels of the two sets should be as large as possible. With a similarity level of at least 320 for similar combinations and of at most 175 for dissimilar combinations, the set for similar combinations contains thirty-nine pairs with an average similarity level of 351 and the set for dissimilar combinations contains forty pairs with an average similarity level of 162. The letter font was Courier News and the size was 20 points.

An example of a similar letter combination is “C” and “O” which has a similarity level of 383. An example of a dissimilar letter combination is “C” and “N” which has a similarity level of 167. The complete list is available at the following URL: https://osf.io/amk5u.

Procedure and Design

In this experiment, there were eight distinct conditions: HH, HL, LH, LL, E_C, I_C, E_S and I_S, which was described in the Introduction. For half of these conditions (HH, HL, LH and LL), the correct response was *Same*, and for the other half (E_C, I_C, E_S and I_S), the correct response was *Different*. Figure 4.1 depicts the eight conditions.

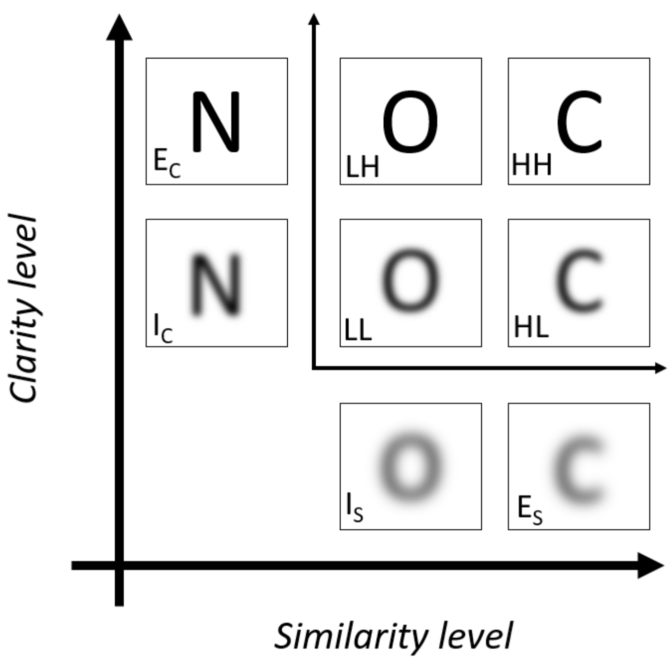


Figure 4.1. Schematic example of the eight conditions. In this example, S1 is a clear “C” and each box is an example of S2 for each condition. The x-axis corresponds to the similarity level of S2 (right = identical, middle = similar, left = dissimilar) and the y-axis corresponds to the clarity level of S2 (top = clear, middle = slightly blurry, bottom = very blurry). For the four top-right conditions, the correct response is Same and for the remaining conditions, the correct response is Different.

Before the experiment, the participants were given instructions about the task and details about what similar and dissimilar letters consist of, as well as examples of slightly blurry and very blurry letters. Additionally, each participant had to complete a practice phase, in which they saw each condition four times, for a total of 32 practice trials.

A trial started with a fixation point located at the centre of the screen for 506 ms. The fixation point disappeared, leaving the screen blank for 400 ms. S_1 appeared above the centre of the screen (visual angle of about 4°), for a duration of 400 ms. The screen was left blank again for 400 ms, and S_2 appeared below the centre of the screen (visual angle of about 4°) until the participant responded or if no response was registered after 5000 ms (in that case, the trial would be considered as having timed out). Participants were instructed to press the *Control* and *Enter* keys to indicate their response as fast and as accurately as possible. During practice trials, participants received feedback on both correct (green text located at the bottom of the screen reading “Correct response”) and incorrect trials (red text located at the bottom of the screen reading “Incorrect response”) for 506 ms. During non-practice trials, feedback was only provided on incorrect trials.

Each condition was replicated 252 times, for a total of 2016 trials. Participants completed the task in two sessions on separate days. The conditions were presented in a random order, but the number of trials per condition was identical in both sessions. The key assignment was counterbalanced across subjects (half used the *Control* key to respond *Same* and the *Enter* key to respond *Different*, and vice versa for the other half). At each quarter of the experimental session, participants were offered breaks. The background of the screen had a luminance of 0.06 cd/m^2 and the stimuli a luminance of 30.80 cd/m^2.

Participant and Power Analysis

As was preregistered, the participants are twenty undergraduate and graduate students. Each participant received $20 CAD for their participation in the study. The experimental protocol was approved by the University of Ottawa's Research Ethics Board.

There is no specific power analysis associated with SFT, mainly because the analyses are based on distribution of RTs. Therefore, statistical power comes from the number of trials, rather than the number of participants. For instance, there were 252 trials in each experimental condition, which represents statistical power well above 95% for the Kolmogorov-Smirnov tests used to analyze SIC(t) curves (Boyerinas, 2016).

That said, we still wanted to recruit twenty participants, as is customary in our laboratory for comparison tasks. Such sample size can detect a fast-same effect about equal to a Cohen's *d* of 0.4, which is typically observed, with statistical power above 90%.

Preregistration and Data Availability

This methodology, including the stimulus manipulation, the procedure, the number of participants and the analysis plan, was preregistered, and the form is available at the following URL: https://osf.io/amk5u. This web page also hosts the materials used in the task, the raw data and the analysis scripts.

Results

In total, twenty participants (11 females, 9 males; aged 18 to 35), all students of the University of Ottawa, were recruited by the experimenter to complete the experiment. They completed 40,320 trials. We removed 105 trials from our analyses for the following reasons. On 67 trials, no response was detected after 5000 ms. Participants responded too rapidly (< 100 ms) on 11 trials, or too slowly (> 4000 ms) on 27 trials. We decided to remove those trials as it is unlikely that participants were realizing the task as instructed on these trials. This represents a removal rate of 0.26%.

We will analyze these trials as follows. First, we explore the RTs and accuracy of participants and examine the effectiveness of our manipulations. Second, we directly test our main hypotheses using an SFT analysis at the group level. Finally, we report individuals for which the results differed from the group. For the analyses using mean RTs analyses, we calculated the mean RT of every participant on correct trials only. For the SIC(t) analysis, we estimated the survival distribution in every condition. The survival distributions were estimated for every participant, but also for the group, using the technique described in Cousineau, Thivierge, Harding, & Lacouture (2016) for averaging individual distributions.

Mean RTs and Accuracy Exploration

Although the comparison task in this study is more complex than what is typically asked from participants, they still responded with good accuracy, the lowest being in the LL condition at 86.4% ± 2.2% and the LH condition at 90% ± 2.0%, where ± is used herein to denote the standard error of the mean across participants. In these conditions, participants had to identify that the two letters were similar. On the other end, the two most accurate conditions were HH (96.5% ± 0.7%) and HL (95.7% ± 0.8%).

Despite the unusual design for a comparison task, participants still responded *Same* faster than *Different* by about 42 ms ± 5 ms, $t(19) = 7.73$, $p < .001$, Cohen’s d =

1.73, BF = 39,164.[12] Looking at mean RTs by condition, we find that this effect is, however, only true when the letters are identical (the HH and HL conditions; fast-same effect of 93 ms ± 7 ms). When the letters are similar (the LH and LL conditions), the *Same* mean RTs are slightly slower than the *Different* conditions (mean difference of 14 ms ± 8 ms), along with having smaller accuracy. This suggests that the fast-same effect is bounded to physical identity and not to the response *Same* itself. That said, if the letters are identical but the response is *Different* (i.e., the E_S condition), the mean RT is about the same as the other *Different* response (581 ms ± 21 ms). This suggests that participants have less information to process in the HH and the HL conditions compared to the six other conditions, which is surprising considering that the task requires an exhaustive processing of both similarity and clarity. We will come back to this result in the discussion.

Table 4.1. Summary of mean response time and accuracy.

Condition	Mean Response Time	Mean Accuracy
HH	489 ± 19 ms	96.5% ± 0.7%
HL	511 ± 22 ms	95.7% ± 0.8%
LH	591 ± 24 ms	90.2% ± 2.0%
LL	617 ± 25 ms	86.4% ± 2.2%
E_C	605 ± 24 ms	93.7% ± 0.9%
I_C	599 ± 25 ms	94.6% ± 0.9%
E_S	581 ± 21 ms	93.3% ± 1.2%
I_S	574 ± 22 ms	95.3% ± 0.7%

Notes. N = 20. Mean are displayed with ± 1 standard error of the mean.

[12] The Bayes Factor is calculated using an informed prior following a Cauchy distribution with a location parameter of 0.4 and a scale parameter of $\sqrt{2}$ (Wagenmakers et al., 2018).

To summarize, fast mean RTs only seems to occur when the second stimulus is identical to the first stimulus, but not when the second stimulus is similar to the first stimulus. Mean RTs and accuracy are given in Table 4.1.

SFT Analyses at the Group Level

The distributions of all twenty participants were averaged to conduct the SFT analysis described in this section. Kolmogorov-Smirnov dominance tests show that participants were faster in the HH and HL conditions compared to the LH and LL conditions. This means that it was easier for participants to establish that the correct response was *Same* when the letters were identical compared to when the letters were similar, as anticipated. However, the HH condition was not much faster than the HL condition, nor was the LH condition much faster than the LL condition (mean differences of 22 ms and 26 ms respectively, within the standard error of the mean which are on average ± 22 ms). This means that the effect of slightly blurring S_2 did not render the task much more difficult, at least from a distribution of response time perspective. We nonetheless proceed with our main analysis.

Mean Interaction Contrast.

The mean interaction contrast (MIC) focuses on how the mean RTs interact in the *Same* conditions. Using a 2×2 repeated-measures ANOVA, we can estimate how the processing of information about similarity and the processing of information about clarity interacted to affect mean RTs. As seen in Figure 4.2, the interaction is negligible (with MIC = -1.10 ms), strongly suggesting that participants processed similarity and clarity serially ($F(1, 19) < 0.04$, $p = .852$).

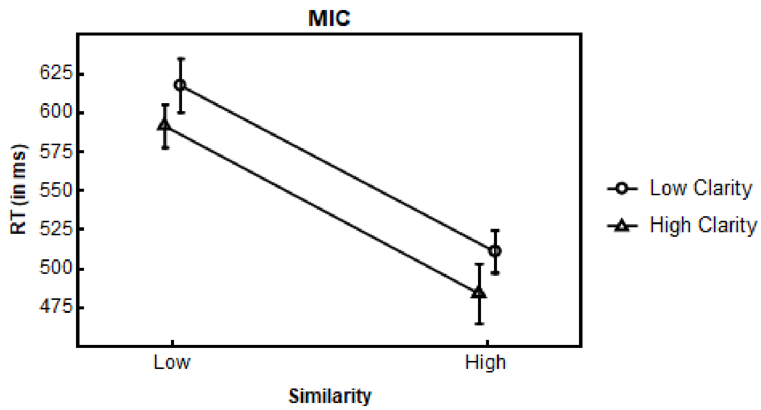


Figure 4.2. Mean response times (in ms) of the four Same conditions depicting the mean interaction contrast (MIC). Error bars are difference- and correlation-adjusted 95% confidence intervals.

Survival Interaction Contrast.

The survival interaction contrast SIC(t) is essentially the same analysis as the MIC, but for every millisecond of the survival functions instead of the means. Each cognitive architecture is characterized by its own prototypical SIC(t) curve (Townsend & Nozawa, 1995). We plotted the group curve in Figure 4.3 (in black), along with the curve for every individual participant (in gray and pink). As seen, the curve tends not to deviate from zero. Using two Kolmogorov-Smirnov tests, we find that the curve's maximum is not different from zero, D = 0.02, p = .970, and that the curve's minimum is not different from zero, D = -0.06, p = .658. Again, this is very typical of a serial architecture.

Unexpectedly, the characteristics of the observed SIC(t) curve suggest a self-terminating stopping rule. Since participants must process both the similarity and the clarity of the letters to respond, we should logically observe an exhaustive stopping rule. This result is difficult to interpret. Perhaps participants do not immediately process the second stimulus in detail. Instead, they try to grasp a more general idea of the characteristics of the stimulus. For example, if the stimulus seems blurry, participants might (not consciously) decide to focus their attention on that dimension instead of similarity. Conversely, if the global shape of the stimulus seems dissimilar, they might decide to focus their attention on the similarity between the two stimuli. In that sense, participants might not process both dimensions before responding. It might be a two-stage process: in the first stage, a pre-attentive scanning of the stimulus guides the search for the second stage, an attention-focus process that verifies the intuition of the first stage. If this interpretation is correct, participants should process similarity and clarity in a mixed order. We test this hypothesis in the following subsection.

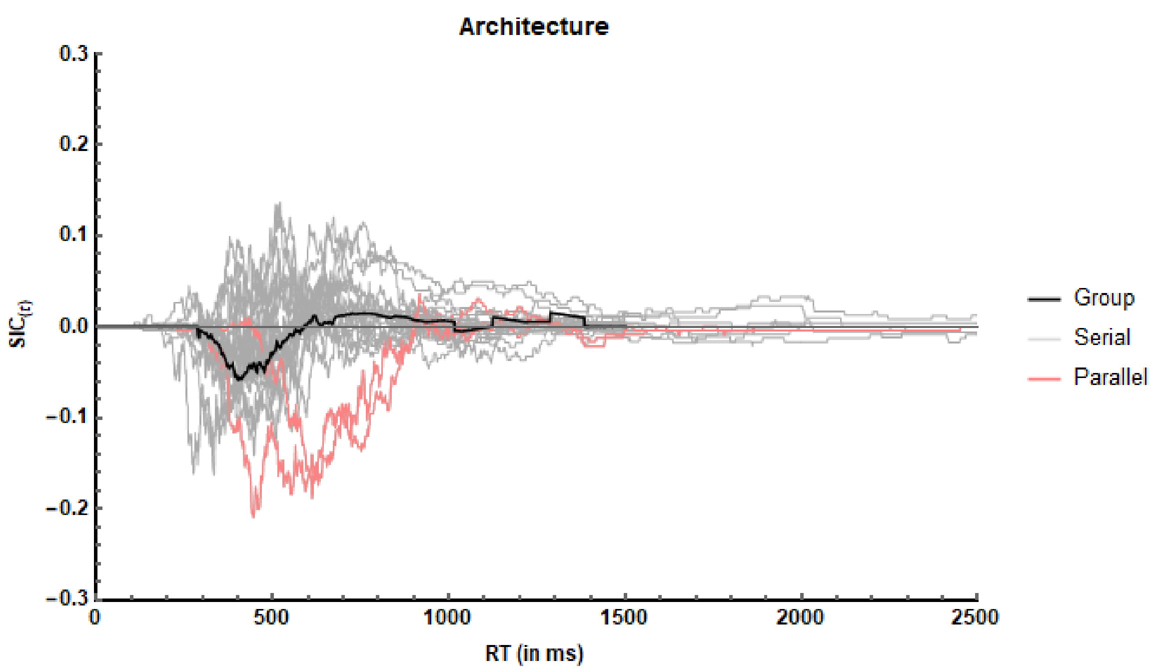


Figure 4.3. Survival interaction contrast SIC(t) curve of the group (in black) and every single participant (in grey for a serial architecture and in pink for a parallel architecture).

Interiors and Exteriors

The MIC and SIC(t) analyses show that participants likely processed similarity and clarity in serial. However, it remains unclear whether the order of treatment was fixed or random. According to the dual-process model, participants should always process information about similarity first, then information about clarity. We can test this hypothesis using the *Different* trials, or the Interiors and Exteriors conditions (I_S, I_C, E_S and E_C). If the order is fixed to similarity-first, clarity-second, we should observe comparable mean RTs in the I_C and E_C conditions, and slower mean RTs in I_S compared to E_S. As seen in Figure 4.4, the mean RTs in all four conditions are very comparable. Using a 2×2 ANOVA, we see that not only the interaction is negligible (MIC = -2 ms; $F(1, 19) = 0.01$, $p = .928$) but the main effect of the difficulty is also quite small (7 ms; $F(1, 19) = 2.62$, $p = .122$). The effect of dimensions, on the other hand, is much stronger (25 ms; $F(1, 19) = 7.36$, $p = .014$). This suggests that participants processed information in a random order, contrary to what is suggested by the dual-process model.

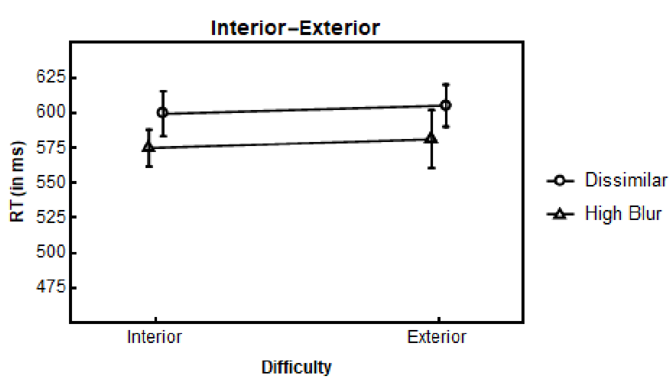


Figure 4.4. Mean response times (in ms) of the four Different conditions depicting the Interior-Exterior interaction. Error bars are difference- and correlation-adjusted 95% confidence intervals.

Table 4.2. Summary of the systems factorial technology analysis for every participant.

Subject	MIC	SIC(t)+	SIC(t)-	Int.-Ext.	Summary
1	-2.94, *p* = .893	0.13, *p* = .115	-0.10, *p* = .321	0.39, *p* = .534	Serial; Random Order
2	20.77, *p* = .269	0.07, *p* = .523	-0.07, *p* = .518	0.10, *p* = .751	Serial; Random Order
3	11.77, *p* = .733	0.06, *p* = .632	-0.06, *p* = .665	0.47, *p* = .494	Serial; Random Order
4	46.87, *p* = .023	0.12, *p* = .159	-0.06, *p* = .633	0.25, *p* = .616	Serial; Random Order
5	-13.31, *p* = .351	0.02, *p* = .949	-0.16, *p* > .999	5.98, *p* = .015	Serial; Similarity First
6	1.55, *p* = .949	0.11, *p* = .213	-0.04, *p* = .846	0.38, *p* = .540	Serial; Random Order
7	-10.06, *p* = .777	0.04, *p* = .855	-0.11, *p* = .223	0.63, *p* = .427	Serial; Random Order
8	31.63, *p* = .523	0.05, p = .744	-0.06, p = .627	0.98, p = .322	Serial; Random Order
9	25.19, *p* = .099	0.12, *p* = .213	-0.04, *p* = .822	3.59, *p* = .059	Serial; Random Order*
10	18.12, *p* = .222	0.10, *p* = .325	-0.14, *p* = .108	1.90, *p* = .168	Serial; Random Order
11	-13.37, *p* = .760	0.11, *p* = .323	-0.13, *p* = .198	0.530, *p* = .467	Serial; Random Order
12	-5.83, *p* = .777	.06, *p* = .633	-0.11, *p* = .225	0.489, *p* = .484	Serial; Random Order
13	-43.32, *p* = .086	0.03, *p* = .885	-0.19, *p* = .016	8.19, *p* = .004	Parallel
14	2.11, *p* = .910	0.08, *p* = .449	-0.14, *p* = .105	0.03, *p* = .856	Serial; Random Order
15	15.72, *p* = .377	0.10, *p* = .289	-0.06, *p* = .618	1.18, *p* = .278	Serial; Random Order
16	-3.49, *p* = .889	0.08, *p* = .480	-0.09, *p* = .395	4.63, *p* = .032	Serial; Similarity First
17	6.45, *p* = .764	0.10, *p* = .300	0.09, *p* = .399	0.06, *p* = .800	Serial; Random Order
18	-3.20, *p* = .846	0.07, *p* = .559	-0.05, *p* = .559	1.79, *p* = .181	Serial; Random Order
19	-44.04, *p* = .398	0.05, *p* = .721	-0.06, *p* = .629	0.01, *p* = 923	Serial; Random Order
20	-58.29, *p* = .058	0.03, *p* = .864	-0.21, *p* = .005	1.178, *p* = .278	Parallel

Notes. Mean interaction contrasts (MIC) tests are a 2×2 analysis of variance for the factors similarity and clarity for the *Same* conditions. Survival interaction contrasts SIC(t) are Kolmogorov-Smirnov tests evaluating if the maximum and if the minimum of the curve differs from zero. Interior-Exterior tests are a 2×2 analysis of variance for the factors similarity and clarity for the *Different* conditions. The asterisk (*) denotes a result that is only slightly above our preregistered cut-off (see details in text).

That said, we must interpret this result with care, as we did expect the difficulty to influence the RT in at least one of the conditions.

Individual Analysis

As planned, we also conducted the SFT analysis for every participant. Table 4.2 provides a summary of the findings. In this section, we discuss participants that deviated from the group, either for their architecture or for their order of treatment.

Participants 13 and 20.

Two participants out of twenty exhibited a parallel architecture. This means that these two participants processed information about similarity and about clarity simultaneously. Their SIC(t) curve is highlighted in pink in Figure 4.3.

Participant 13 is also characterized by unorthodox mean RTs. This participant responded, on average, in 700 ms in both *Same* and *Different* trials. This is the only participant for which we do not observe a fast-same effect. That said, this participant was still much faster in the HH condition compared to any other conditions, averaging a RT of 609 ms. The most likely scenario is that this participant relied heavily on perfectly matching stimulus to accumulate evidence in favour of the *Same* response and struggled to gather *Same* evidence when the stimuli were similar or slightly blurry, that is compared to other participants. In contrast, Participant 20 responded, on average, *Same* in 609 ms and *Different* in 697 ms.

Participants 4, 9 and 16

Some participants that exhibited a serial architecture also had a fixed order of treatment, processing information about similarity first. Whereas the effect is quite strong for participants 4 and 16, it is marginally above the preregistered cut-off threshold of $\alpha = .05$

for participant 9 (*p* = .059). These results suggest that two or three participants out of twenty processed information about similarity in priority over the information about clarity, fully supporting the dual-process model. However, these participants did not seem to benefit more from a fixed-order processing compared to the other participants that had a random order, in the sense that their fast-same effect (HH and HL condition compared to *Different* conditions) are comparable (respective raw fast-same effect of 88 ms, 55 ms and 79 ms).

Despite observing individual differences in terms of architecture and order of treatment, we observed the fast-same effect in almost all the participants, except for participant 13. This suggests that serial processing with priority for similarity is not causally related to the fast responses observed in *Same* trials.

Discussion

The main objectives of this study were to identify the cognitive architecture underlying the comparison task and if this architecture is serial, to determine if the order of treatment was fixed or random. These research questions were ingrained in the dual-process model, which postulates that the architecture is serial and that the order of treatment is fixed (similarity is processed first). Our results suggest that the architecture is indeed serial, as hypothesized. However, they also suggest that the order of treatment is random, contrary to the dual-process model assumption. They also suggest that participants use a self-terminating stopping-rule, which is surprising given that the task is supposed to be exhaustive. Additionally, individual differences regarding architecture and order of treatment do not influence the fast-same effect. In other words, cognitive architecture may not be causally related to the fast-same effect.

As a first approximation, researchers were right to discard the dual-process model. However, the dual-process model was correct regarding the seriality underlying the comparison task. Indeed, it seems that participants process information one after the other, unlike what is assumed by alternative models (i.e., the facilitation model of Proctor, 1981; the drift-diffusion model of Ratcliff, 1981).

Performance on *Similar* Trials

A key result of this study is that the fast-same effect seems to disappear in the *Same-Similar* trials (conditions LH and LL). The mean RTs of participants in these conditions were comparable to those in the *Different* conditions. Also, participants were less accurate in those conditions. Evidently, the identification of *similar* letter asked more to the participants, but we did not expect that it would cancel the fast-same effect. Typically, we can still observe an RT advantage even if *Same* stimuli are not physically matching (Posner & Mitchell, 1967; Proctor, 1981, 1986). For example, the fast-same effect is still observed when participants need to match uppercase to lowercase letters, or even at a more categorical level, when participants need to match vowels together, or consonants together. Our results suggest that this does not extend to task in which letters are physically matching in some conditions.

A potential explanation for this result is that, just like in *Different* trials, participants do not know what to expect in terms of a *similar* stimulus. Say that S_1 is “C”. An identical letter is easy to represent for the participant, it is a “C”. However, what is a dissimilar “C”? It could be a “F” or a “N”; the participants do not really have a mental representation of what a dissimilar “C” looks like. *Similar* trials are also characterized by a lack of a mental representation. What is a similar “C”? It could be a “O” or a “G”.

When S_2 is identical to S_1, participants realize a one-to-one mapping of their mental representation of the stimulus to the displayed stimulus to provide a correct *Same* response. However, when S_2 is similar or dissimilar to S_1, participants need to realize a many-to-one mapping to respond correctly (many possible alternatives to one displayed stimulus). Farell (1985) argued that many-to-one mapping requires participants to use more attentional resources, which could explain why the response time is longer.

The model of Farell states that participant first realize easier tasks (one-to-one mapping), because these tasks require fewer resources. If the one-to-one mapping is inconclusive, then participants must use attentional resources to analyze the stimulus more thoroughly. In that sense, this model is highly similar to the dual-process model, but proposes that participant only use a single processor to realize the task (not two distinct) and that the processor can either respond quickly using a one-to-one mapping strategy, or use attentional resources for a many-to-one mapping, at the cost of extra time. Such model would explain why, in our study, *similar* trials were as slow as *Different* trials.

Regarding our results, it would also explain why we observed a self-terminating stopping-rule in a seemingly exhaustive task. If participants perform by default a pre-attentive scan of the stimulus, they also likely gather some initial intuition about what the stimulus globally looks like. This intuition probably guides the analytical process afterwards. For instance, if the pre-attentive process observes a globally high level of blurriness, it will likely focus the attention of the participant towards processing clarity first. A combination of the pre-attentive intuition and the attention-driven processing might result in the observed self-terminating stopping-rule. The order of treatment is not completely random, it is influenced by a pre-attentive scan of the stimulus (Cousineau & Shiffrin, 2004).

Limitations of the Current Approach

The unexpected slow *similar* trials represent a limitation of the current experimental design. Potentially, participants sequentially ponder three response alternatives for similarity: is S_2 identical, similar or dissimilar to S_1? instead of merging identical and similar stimuli together in a single alternative. If such is the case, *similar* trials were not simply a weaker signal for the *Same* response, but another treatment in itself. This means that maybe the order of treatment is fixed, but that our design was not able to detect it. Nevertheless, we think that our study helped better understand the cognitive mechanisms underlying a *Same-Different* task. We now know that participants do not prioritize similarity, rather simplicity. This finding contradicts the proposition of some researchers that participants have an inherent bias for similarity (Irwin et al., 2001; Ratcliff, 1978; Ratcliff & Hacker, 1981; Ratcliff et al., 1989; D. A. Taylor, 1977).

An important result of our study is that we observe individual differences, both quantitatively and qualitatively. Five participants out of the twenty recruited had either a different architecture or a different order of treatment than the rest. The effects of the manipulations on RTs also varied largely across participants. For many participants, clarity was much easier to detect than similarity. This is evident when comparing the main effect of the similarity manipulation to the main effect of the clarity manipulation. Fifić and Little (2017) refer to this phenomenon as *unequal stretching*. A potential solution for this problem would be to adapt the intensity of the manipulation on a participant basis (Houpt & Fifić, 2017). That said, whereas this would be easy to implement for the clarity manipulation (albeit it would increase the number of sessions), it would be much harder to accomplish for the similarity of the letters.

Another potential limitation of our design is that we used letters as our stimuli. We decided to use letters for multiple reasons. First, it is the most commonly used stimulus in comparison task studies. Second, it renders *similarity* easy to manipulate (using indices such as similarity levels to create our stimuli). Third, participants are already very familiar with the stimuli. Using stimuli unknown to participants could have considerably increased the difficulty of an already complex task (compared to more classical versions of comparison tasks). That said, some might argue that letters provide additional information (i.e., phonological information) that participants need to inhibit in order to realize the task accurately. However, it is difficult to find stimuli that do not add irrelevant noise or redundant signal to the treatment. Participants label stimuli all the time to facilitate their treatment and recall them more easily in working memory.

Final Remarks and Future Directions

The present study is the first attempt at directly testing the cognitive architecture assumption underlying the dual-process model of the comparison task. This model predicted that participants would process information about similarity in priority, because of the existence of an identity processor. Our results suggest that this is not the case, and that participants seem to rather prioritize simpler treatment of information, and not similarity. That said, most participants exhibited a serial treatment, as predicted by the dual-process model. Regardless, we propose to abandon this model.

Considering limitations of this study, future research should focus on investigating in more depth the role of attention in comparison tasks. Using cues to direct attention, researchers could create a one-to-one mapping of a *Different* response. For example, an auditive cue could be associated with a specific letter. Participants could therefore keep in working memory a second reference letter (the *Different* letter). Another avenue would be to use non-letter stimuli. The advantage would be to extend the current results to more generalizable findings. That said, varying the similarity level of a non-letter stimulus can be difficult.

A final direction would be to investigate the cognitive architecture involved when identifying the attributes composing the stimulus. In the current study, we were interested in the architecture needed to process distinct response dimensions (similarity and clarity), but it would also be important to assess how participant process information within the stimulus itself. To that end, we already develop an experiment, for which the design has been approved as a Registered Report (Goulet & Cousineau, 2020). This study will also explore how additional information affect the efficiency of the treatment (the processing capacity). If participants indeed require more attentional resources to respond *Different* compared to *Same*, they should have better processing capacity in the latter than the former.

Chapter 5 – Study 3: Cognitive Architecture and Capacity of the Cognitive System Responsible for *Same – Different* Judgments

Abstract

Participants tend to match identical pairs of stimuli faster than different pairs. Despite many endeavours to explain this fast-same effect, there is still no theoretical consensus. A potential reason for the lack of consensus is that the cognitive architecture and capacity underlying such phenomenon is assumed and not formally tested. For example, the dual-process approach suggests that Same responses arise from a parallel treatment, whereas Different responses arise from a serial treatment. It also suggests that in both conditions, the capacity of the process is unaffected by workload (unlimited capacity). Alternative approaches argue that the fast-same effect can be explained by parallel or coactive architectures with channels working in either limited or super capacity. In this study, we formally assess the architecture (three possibilities: serial, parallel and coactive) and the capacity (three possibilities: unlimited, limited and super-capacity) of the cognitive system in a Same-Different task using Systems Factorial Technology (SFT). We recruited twenty participants to perform a double-factorial task lasting four sessions. Because of the lack of effectiveness of the blurring manipulation, we cannot draw a strong conclusion about the cognitive architecture. As for the capacity, the results show that it is mostly limited for the majority of participants. However, between 300 and 500 ms, participants tend to have a much stronger processing capacity in the Same condition compared to the Different condition. This short but strong burst of activity for identical stimuli might explain the fast-same effect.

Reference of the article

Goulet, M.A., & Cousineau, D. (2020). Cognitive Architecture and Capacity of the Cognitive System Responsible for Same – Different Judgments. *Attention, Perception, & Psychophysics*. Advance online publication. https://doi.org/10.3758/s13414-020-02008-z

Introduction

The assessment of similarity and differences in objects is a basic cognitive behaviour performed daily. Typically, this ability is studied by presenting pairs of stimuli to participants whose task is to indicate if the stimuli are *Same* or *Different*. One intriguing result of such a task is that participants respond quicker in *Same* trials compared to *Different* trials (e.g., Bamber, 1969; Egeth, 1966; Nickerson, 1965; Posner & Mitchell, 1967). This result contradicts a first intuitive theory to explain similarity judgment in which identifying a *Same* pair requires an exhaustive treatment of all the stimulus's attributes. On the other hand, a *Different* pair only requires the identification of one mismatching attribute. Therefore, this theory posits that participants would respond faster in the *Different* trials.

This theoretical approach has its origin in the visual search literature. In this task, participants must identify a target among distractors (Holmgren et al., 1974). As soon as the participant identifies the target, they have enough information to respond, but in trials without targets, the search must be exhaustive and process all the distractors before concluding that there is no target. This strategy was explored in the Same-Different task with what was dubbed *disjunctive tasks* in which participants are asked to detect the presence of any matching attribute between two stimuli (Bamber et al., 1975; Derks, 1972; Farell, 1977; Nickerson, 1967a; Sekuler & Abrams, 1968; Silverman & Goldberg, 1975; D. A. Taylor, 1976a). In the disjunctive task, participant must find at least one matching attribute to respond *Same*. The results of such task show an inversely proportional relationship between the number of critical attributes (the number of matches) and the response time. Symmetrically, in the classical versions of the *Same*-

Different task (the *conjunctive task*), participant must detect at least one mismatching attribute to respond *Different* (Farell, 1985; Hyun et al., 2009). If the symmetry was perfect, one would expect that in the conjunctive task, the slowest condition would be the *Same* condition (the one with zero critical attributes). However, in such task, the opposite is observed: the *Same* condition is the fastest.

Many experiments explored the fast-same effect by manipulating the stimuli (mainly letters and geometric shapes; see Bamber, 1969; Nickerson, 1967a, 1967b; but also line lengths, tones and motion directions; see Bindra, Donderi, & Nishisato, 1968; Bindra, Williams, & Wise, 1965; Link & Tindall, 1971; Petrov, 2009) and the design / instructions (Proctor & Rao, 1983b; Proctor et al., 1984; Ratcliff & Hacker, 1981; D. A. Taylor, 1977). Yet, as spelled out in the many review articles on the effect, there is no consensus on why this phenomenon occurs (Farell, 1985; Krueger, 1978; Luce, 1986; Nickerson, 1973b, 1978; Ratcliff, 1985; Sternberg, 1998).

A main reason for the lack of consensus if that these theoretical explanations are not, in most cases, framed in terms of the cognitive architecture and capacity of the cognitive system underlying the task. Most theoretical explorations focused on questions regarding variations on decision thresholds, encoding rates and evidence accumulation rates. However, they disregarded how the process of evidence accumulation is implemented. Indeed, most investigations on the fast-same effect occurred prior to the publication of Townsend’s methodology for diagnosing cognitive architecture and capacity (Townsend & Ashby, 1983; Townsend & Nozawa, 1995). This *taxonomy of elementary cognitive processes* was not ignored; most of the explanations supposed a specific organization of the processing channels. However, the architecture underlying these theories were not formally tested, simply assumed.

The objective of this study is twofold: to formally assess 1) the cognitive architecture and 2) the capacity of the cognitive system responsible for *Same* and *Different* judgments. We focus on how evidences are accumulated (i.e., one at a time, in parallel or collaboratively) and how the cognitive processes respond to an increase in the total amount of information presented (i.e., are they unaffected, hindered or facilitated?). Although these research questions do not directly address potential variations in thresholds, encoding rates and evidence accumulation rates, it will provide a more solid framework to model these parameters.

Cognitive Architectures

In a seminal paper on the *Same-Different* task, Bamber (1969) proposed that two processes are involved in the response. The first processor is an *identity reporter* which processes a chunked version of the stimulus to detect solely similarity. The processing units (or *channels*) of the identity reporter are assumed to be organized in parallel so that all the attributes are processed simultaneously. The second processor, responsible for *Different* responses, is analytical: it partitions the stimulus in many attributes for a thorough treatment (Nickerson, 1965). This slower treatment of the attributes would be more apt at distinguishing a mismatch from a match in the presence of internal noise. As mentioned by Krueger (1978), it is more likely that noise changes a matching attribute into a mismatching attribute than the opposite. For example, a participant is presented with a “J” followed by a “T”. Due to internal noise, it is possible that the second stimulus is truly a “J” that has been perceived as a “T”. Therefore, the participant should process this stimulus more thoroughly to confidently say that it is mismatching. Conversely, if the participant perceives the second stimulus as a “J”, because the probability that a mismatching letter is transformed into a “J” is low, the participant can confidently say that it is matching. For this reason, participants might process mismatching attributes more thoroughly to ensure accurate responses. It is also assumed in Bamber (1969) that the comparisons are performed serially. The attributes are processed one by one and as soon as a mismatching attribute is detected, the treatment is terminated. This conceptualization adequately accounts for the decrease in response time (RT) when the number of mismatches increases (Bamber, 1969; Egeth, 1966; Nickerson, 1967b).

Alternatively, some researchers have argued that *Different* RTs can also be modelled with a parallel architecture (Hawkins, 1969; Hawkins & Shigley, 1972). These authors showed that if the channels have different processing rates, a parallel model can fit the relationship between RT and the number of mismatches. Studies manipulating the position of the mismatching attributes show evidence for discrepancies in channels processing rates, where attributes located at the extremities of the stimulus are processed faster than attributes located at the centre (Howell & Stockdale, 1975; Koriat & Norman, 1989a, 1989b; Proctor, Healy, & Van Zandt, 1991).

Serial and parallel architectures posit that each channel can trigger responses independently of the other channels. On the other hand, the coactive architecture, proposed by Miller (1978, 1982), postulates that the role of the channels is to feed their activation to a single counting unit (an accumulator) which is responsible for responding. Therefore, the channels are working in collaboration: they pool their evidence until the counting unit’s decision threshold is reached for either response alternative.

As seen, the assumed underlying architecture changes how one theoretically explains the fast-same effect. When the serial architecture is assumed, the fast-same effect seems to be the result of a criterion shift, that is, to be caused by discrepancies in decision thresholds. On the other hand, when the parallel architecture is assumed, the fast-same effect seems to be the results of varying evidence accumulation rates. Finally, when a coactive architecture is assumed, the fast-same effect seems to be the results of the architecture itself, and not caused by variations of specific theoretical parameters.

Despite its central role, architectural questions have been omitted by a few more models that posit single-process explanations. These models focused mainly on the accumulation of information over time. They proposed that the fast-same effect can be the result of encoding facilitation (C. W. Eriksen et al., 1982; Krueger, 1978, 1983; Proctor, 1981, 1986; Proctor & Rao, 1983a; St. James & Eriksen, 1992) or response bias for *Same* (Irwin et al., 2001; Ratcliff, 1978; Ratcliff & Hacker, 1981; Ratcliff et al., 1989; D. A. Taylor, 1977). These models likely assumed that the information is gathered in parallel, and that the number of attributes contained in the stimuli has no effect on the processing efficiency (this is also referred to as unlimited capacity, independent parallel – UCIP).

Capacity of the System

Like cognitive architecture, the capacity of the cognitive system has never been tested directly in the *Same-Different* task. Capacity relates to the efficiency of the treatment when the number of attributes – the workload – is increased (Hylan, 1903). In such scenario, the processing efficiency might remain unchanged (*unlimited capacity*) but might also be hindered (*limited capacity*) or augmented (*super capacity*). As put by

Eidels, Houpt, Altieri, Pei, and Townsend (2011), the capacity of a system characterizes whether additional information facilitates or inhibits the system irrespective of the number of processes assumed.

The most commonly found assumption is that the efficiency of the cognitive system is unlimited – that the workload does not facilitate nor inhibit the processing time. This assumption has not been directly tested within the Same-Different task. Assuming by default that the capacity is unlimited can lead to important theoretical misconceptions. For example, a common result of the *Same-Different* task is that increasing the number of attributes slows down the response time (Atkinson et al., 1969; Bamber, 1969; Sternberg, 1966). If processing capacity is unaffected by an increased number of attributes, the slowdown in response time would likely be caused by greater decision thresholds (as more information needs to be processed). However, if the capacity is affected, the discrepancy in response time can be explained by differences in processing rate (Donkin et al., 2014).

A few authors proposed that the capacity of the channels might indeed be limited in the *Same-Different* task (Decker, 1974; Hawkins & Shigley, 1972; Link & Tindall, 1971). However, their proposal was deduced from comparing mean RTs across conditions. Mean RTs can lead to erroneous assumptions regarding cognitive architecture (Godwin et al., 2015; Little, Eidels, Houpt & Yang, 2017; Townsend, 1972). For example, mean RTs cannot distinguish between the time required to perform subtasks in a serial architecture and the time required by each processing channels in a parallel architecture (*intercompletion times* – Townsend, 1972). For this reason, Townsend and Nozawa (1995), among others, suggested using a factorial design and comparing distributions of RTs across conditions instead. Similarly, distributions can be used to gather information regarding the processing capacity of the channels (i.e., Grice, Canham, & Gwynne, 1984; Miller, 1982).

The aforementioned work of Miller (1978, 1982) used distributions of RTs to support not only the coactive architecture but also the super-capacity workload of the system. This is, however, the only study supporting super-capacity in the Same-Different task.

Formal Assessment of Architecture and Capacity

Architecture and capacity questions were assumed in past theoretical endeavours to explain the fast-same phenomenon. About any architecture and any capacity have been assumed (see summary in Table 5.1). The proposed study is therefore an opportunity to falsify a large number of candidate models. In this study, we formally assess the architecture and the capacity of the cognitive system in the *Same-Different* task.

Table 5.1. Overview of the assumed architecture and capacity in selected papers.

Architecture	
Parallel	Hawkins (1969); Koriat & Norman (1989a, 1989b); Proctor et al. (1991); Ratcliff & Hacker (1981) and many more.
Serial	Bamber (1969); Krueger (1978); Nickerson (1965)
Coactive	Miller (1978, 1982)
Capacity	
Unlimited	Assumed by almost everyone.
Limited	Decker (1974); Hawkins & Shigley (1972); Link & Tindall (1971)
Super-capacity	Miller (1978, 1982)

These two research questions can be directly addressed using Systems Factorial Technology (SFT; Townsend & Nozawa, 1995; but also see Harding et al., 2016; Little, Altieri, Fifić, & Yang, 2017). It combines Donders (1969)'s subtractive and Sternberg (1969)'s additive methods to derive architectural information using the entire distribution of RTs instead of the mean RT.

SFT requires a double-factorial design, that is, an experimental manipulation involving two attributes. In a seminal example, the participants were asked to detect the presence of any dot on a computer display (Townsend & Nozawa, 1995). The dot can be on the left part of the screen (attribute 1) or on the right part of the screen (attribute 2). Participants were asked to press an answer key if at least one dot was presented on the screen, and another answer key if no dot was presented. Each dot can either be clear (high perceptibility - H) or barely visible (low perceptibility - L). The four possible experimental conditions (HH, HL, LH and LL) are then used to diagnose the architecture of the processing channels. The first manipulated factor concerns the number of attributes displayed on the screen. The second manipulated factor is the saliency of the attributes.

In order to diagnose the capacity of the cognitive system, the double-factorial design is expanded with four additional conditions, in which only one attribute is displayed on the screen (HX, LX, XH and XL, where X denotes a missing attribute). It is technically possible to have a ninth condition in which no stimulus is displayed on the screen. However, since it does not directly address our research questions, we omit this empty condition in the current study (Houpt et al., 2014).

SFT has three assumptions. First, the manipulation must have a *selective influence* on the attributes. In the Townsend and Nozawa's (1995) dot experiment, the clarity of the leftmost dot must not influence the clarity of the rightmost dot. This assumption is critical to interpret interaction contrasts (Sternberg, 1969). In the present study, we used the *mean interaction contrast* (MIC) and the *survivor interaction contrast* (SIC) described in Appendix A.

Second, to interpret capacity, we must assume *context invariance* whereby the processing rate of each processing channel is not influenced by the total number of attributes in a condition (Ashby & Townsend, 1986; Houpt et al., 2014; Townsend & Eidels, 2011; Townsend et al., 2012; Yang et al., 2018). In more technical terms, the mean (or the variance) processing competition time of a channel processing a single attribute must not differ between conditions in which only this attribute is present, and condition in which other attributes are also present. Context invariance is required to interpret the capacity coefficient of the system.

Third, SFT postulate that the channels responsible for processing information are stochastically independent (Heathcote, Brown, & Cousineau, 2004; Townsend & Nozawa, 1995). In other words, the processing rate of a channel must not be correlated with the processing rate of other channels. There exists stochastic dependency when the system allows cross-talk between the processing channels (Mordkoff & Yantis, 1991). Although the vast majority of models of the Same-Different task posit stochastic independence, cross-talk has been proposed as a potential explanation for the fast-same phenomenon (Eriksen et al., 1982). Recent work has shown that in cases of stochastic dependency, SFT can still be used to assess the type of channel dependency (inhibitory, facilitatory or both – Eidels et al., 2011).

With this expanded double-factorial design, we can compute the MIC, which distinguishes serial architectures from parallel and coactive architectures, and the SIC, which distinguishes between all three architectures, and the *capacity coefficient*, which can assess processing capacity. Appendix A expands on these measures and their interpretability.

In the current study, we use the expanded double-factorial design and SFT to directly assess our two main research questions. First, we diagnose the architecture of cognitive system responsible to make *Same* and *Different* judgments. There are three possible outcomes: the architecture could be serial, parallel or coactive. Second, we assess the effect of the workload on the capacity of the system. Again, there are three possible outcomes: the capacity could be unlimited, limited or super-capacity.

Methods

The experiment consists in a mix between a classical Same-Different Task design (e. g., Bamber, 1969) and Townsend & Nozawa (1995)'s double-factorial design, expanded as per Houpt et al. (2014). Instead of a target present / target absent task, as in Townsend & Nozawa (1995), participants need to indicate whether two arrays composed of one or two letters presented in succession are *Same* or *Different*. The leftmost and the rightmost letters of the second stimulus of the pair are subject to the clarity manipulation of the double-factorial design. In conditions where only one letter is displayed, the participants are asked to ignore the missing attribute. In addition to the clarity manipulation, we also manipulate the identity of the arrays, that is whether they are *Same* or *Different*. In *Same* trials, all the attributes are matching, whereas in *Different* trials, all the attributes are mismatching.

Participants and Power Analysis

The participants are twenty undergraduate and graduate students of the University of Ottawa recruited by the experimenter. Upon the completion of the task, the participants received $50 CAD. This experimental protocol was approved by the University of Ottawa's Research Ethics Board.

Although the SFT analyses are generally performed at the individual level, we still wanted a representative sample. Moreover, this sample size is the usual number of participants recruited in our laboratory for *Same-Different* studies, as it yields a statistical power to detect the fast-same effect (estimated to be a Cohen's d of about 0.4) well above 80% (Goulet & Cousineau, 2019b). Therefore, we stopped data collection at twenty participants.

That said, since our main hypotheses are tested using SFT, the true statistical power of our study resides in the number of trials in each condition for every participant. We opted for 256 trials per condition and for the two responses, or 4096 trials in total, which fits well within four sessions. Therefore, we calculated, at an individual level, the statistical power of 256 replications to detect small effects. For z tests (used for MIC and capacity), it represents a power of 89%, whereas for the Kolmogorov-Smirnov tests (used for SIC), it represents a power well above 95% (Boyerinas, 2016).

Procedure

The stimuli are displayed on a CRT monitor (1024×768 pixels) with a refresh rate of 85 Hz situated at approximately 50 cm of the participant. A trial starts with a fixation point located at the centre of the screen, for a duration of 506 ms. Following the fixation point, the first array of letters – S_1 – (horizontal visual angle of about 5.2°) is presented slightly above the fixation point (visual angle of about 4.0°) for 400 ms, then disappears. The inter-stimulus interval lasts 400 ms, during which a blank screen is presented to the participant. Finally, the second array of letters – S_2 – appears slightly below the centre of the screen (visual angle of about -4.0°) for 5000 ms, or until the participant provides a response.

In all the trials, S_1 is clearly visible and composed of two letters. The experimental manipulation is only applied to S_2. The manipulation consists in altering the clarity of the leftmost letter (attribute #1) and/or the rightmost letter (attribute #2). In conditions where only one letter is presented in S_2, the attribute presented conserve its physical position in the array. In other words, the letter is either displayed on the leftmost or on the rightmost location of the array, and not in the centre.

The task of the participant is to compare the letter(s) displayed in S_2 with the letter(s) displayed in S_1 at their corresponding positions. Therefore, they can perform a maximum of two comparisons: the leftmost letter presented at S_1 with the leftmost letter presented at S_2, and/or the rightmost letter presented at S_1 with the rightmost letter presented at S_2. The participant must not make cross-position comparisons (i.e., the leftmost letter of S_1 and the rightmost letter of S_2, and vice versa).

Stimuli

The stimuli are arrays of one or two letters composed of a random combination of uppercase consonants sampled from a set of twelve (B, C, D, F, J, K, L, N, S, T, V, and Z) taken from Bamber (1969). They are presented in white on a black background using an Arial font. In a single array, no letter is repeated. The participant's task is to indicate if all the attributes (letters) match or if at least one of the attribute mismatches. In the former case, the correct answer is *Same*, and in the latter case, the correct answer is *Different*. A matching attribute signifies that the same letter at the same position is present in both pairs. A mismatching attribute signifies that a letter differs between the two pairs at a certain location.

In some conditions, one or two letters are slightly blurred using the *Blur* function in Mathematica (version 11.2.0.0) with a radial blur of 3 pixels. The participants are asked to ignore this blur. The complete list of stimuli is available at the following link: https://osf.io/cwm26.

Design

In this experiment, four independent variables are manipulated: i) the clarity of the leftmost letter (2 levels: clear or blurry); ii) the clarity of the rightmost letter (2 levels: clear or blurry); iii) the number of displayed attributes at S_2 (2 levels: one or two) and iv) the identity of the stimulus (2 levels: *Same* or *Different*). For trials with a single letter, the letter was on the right or on the left an equal number of times. In total, sixteen conditions are presented an equal number of times to the participants (all possible combination of independent variables). A summary of the conditions is illustrated in Figure 5.1 for a hypothetical trial in which S_1 is composed of the letters "J C". The dependent variable is the RT; it is measured with a precision of ± 1 ms using E-Prime 2.0 (Psychology Software Tools, 2019).

To ensure adequate estimations of the RT distributions in each condition, participants completed a total of 4096 trials, 256 trials per condition. Because this number represents nearly four hours of testing per participant, the experimental task is divided in four sessions of 1024 trials each. A single session contained 64 trials of each condition. Within a single session, the conditions are randomized. Consequently, on any given trial, the participant could not predict the condition.

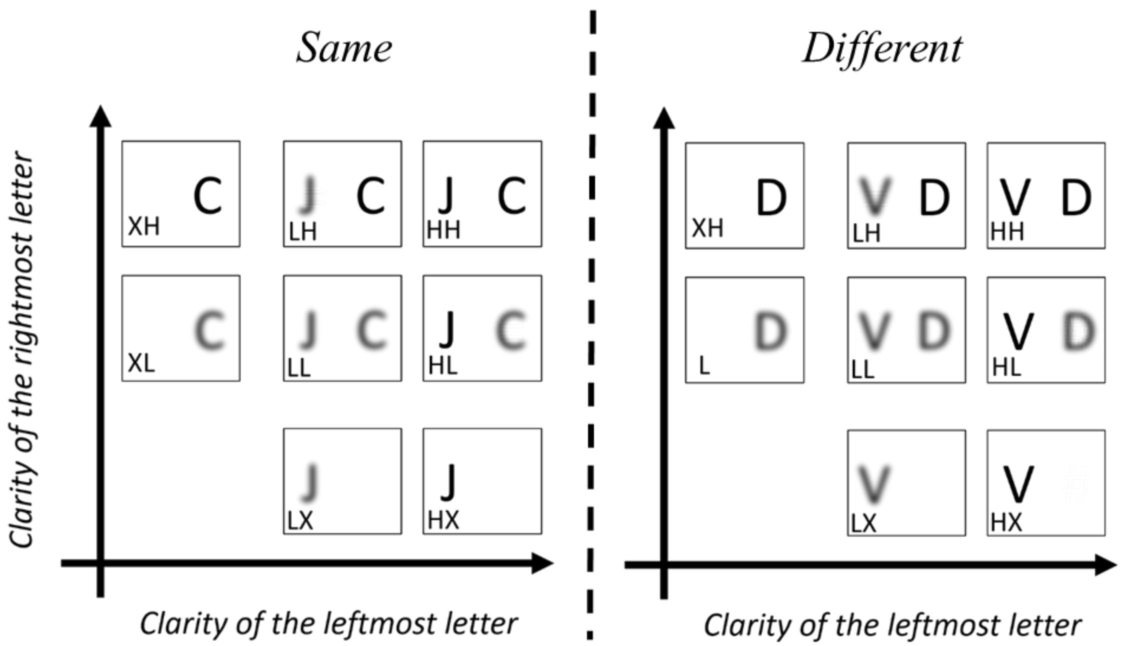


Figure 5.1. Experimental design of the current study. In this example, S1 (not shown) is composed of the letters "J C". Each box represents a valid S2 for all 16 possible conditions. On the left plot, the two letters are matching, whereas on the right plot, the two letters are mismatching. The four top-right conditions of the left plot will be used to assess the architecture of the system responsible to make Same judgments, whereas those conditions in the right plot will be used to assess the architecture with regards to Different judgments. The other conditions will be useful to assess the capacity coefficient of the system.

Data Availability

The raw data (before outlier removal) and analysis scripts are available on the Open Science Framework (OSF) page for this project: https://osf.io/cwm26.

Results

All the analyses reported in this section were part of our preregistered analysis plan. We present the results of these analyses in four sections. In the first section, we discuss the outlier removal and data cleaning. In the second section, we assess the effectiveness of our manipulations. Precisely, we compare mean RTs for *Same* and *Different* trials across participants, and compare distributions of RTs in the HH, HL, LH and LL conditions. In the third section, we address the question of cognitive architecture and report the mean interaction contrast and the survival interaction contrast for the group data and for individual participants. Finally, in the fourth section, we address the question of processing capacity.

Outlier Removal and Data Aggregation

As established prior to data collection, we removed any trial that timed out (no response after 5000 ms), that was too slow (above 4000 ms) or that was too fast (below 100 ms), as it is unlikely that participants were doing the task as requested for these trials. In total, 172 trials were removed out of 81,920, corresponding of a removal rate of 0.21% of the total trials.

Following our preregistered exclusion criteria, we excluded the data of one participant (#12) from the group analyses because their overall accuracy was three standard deviations below the mean of all participants (*z* score of -3.39). However, the data of this participant was still analyzed at an individual level.

For group analyses, mean RTs represent the mean across all participants. The “average distribution” of RTs across participants was obtained using the technique described in Cousineau, Thivierge, Harding, and Lacouture (2016). This technique consists in aggregating 1 ms bins by computing the geometric mean across participants for every bin after removing the estimated base response time of individual participants (T_0) and adding the average T_0 to the computed averaged quantiles (also see Thomas & Ross, 1980).

Fast-Same Effect and Manipulation Effectiveness

Before testing our main hypotheses, we verify that 1) the fast-same effect is observed and 2) our manipulations are effective. For the former, we looked at mean RTs in every condition (HH, HL, LH, LL, HX, LX, XH and XL) and for each response (*Same* and *Different*). We originally planned to conduct a paired-sample t-test between the mean RTs in the *Same* trials and the mean RTs in the *Different* trials. However, after plotting the mean RTs (see Figure 5.2), we decided to change our analysis plan to focus on RT differences between *Same* and *Different* trials by separating conditions in which two letters were presented at S_2 (HH, HL, LH and LL) and conditions in which one letter was presented at S_2 (HX, LX, XH and XL).

In trials for which two letters were presented at S_2, the fast-same effect is about 32 ms, 95% confidence interval, CI, of [15,50], $t(18) = 3.97$, $p < .001$, BF = 24.38, $d = 0.91$.[13] This corresponds to what is typically observed in *Same-Different* tasks. However, we do not observe the same trend for trials in which only one letter is presented at S_2. To investigate the RTs, we used a 2×4 repeated-measures analysis of variance (ANOVA) on the factors *Response* (*Same* or *Different*) and *Condition* (HX, LX, XH and XL). We

---

[13] For the Bayes factor, our prior is a Cauchy distribution with a location parameter = 0.4 and a scale parameter = $(\sqrt{2}/2)$ following the recommendation of Wagenmakers et al. (2018).

found that these two factors interacted (F(3,54) = 4.214, p = .009, BF = 33.59, η_p^2 = .190).[14] Indeed, there seems to be two elements that affect RT. First, when the letter is blurry, the mean RT increases in both *Same* and *Different* conditions. This trend was expected and shows that the blurring manipulation was effective. Second, when only the leftmost letter was presented, participants were faster to respond *Same* compared to when only the rightmost letter was presented. This effect was not observed when participants were responding *Different*. A consequence of this is that the fast-same effect is reduced in trials in which only one letter is presented at S_2 and even reversed when it is the leftmost letter that is absent.

Typically, when a single letter is presented in both S_1 and S_2, the fast-same effect is still observed. In our experiment, when only one letter was presented at S_2, there were two letters presented at S_1. Therefore, in those conditions in our study, S_2 does not perfectly match S_1 (there is one letter missing). This suggests that even though participants only need one matching letter to respond *Same*, they still expect to encode two letters, and set their decision threshold accordingly. If such is the case, participants would lack activation in these trials, resulting in slower RTs. Yet, this does not explain the differences for HX/LX vs. XH/XL trials. It might be due to a serial treatment of the information, which we assess in our next planned analyses. The question remains why such discrepancy is not observed in *Different* trials. We will discuss this strange result later in the Discussion.

---

[14] Because this analysis was not planned, we used a null prior.

Another potential explanation for the reduced or reversed fast-same effect in single-letter trials is response competition. In these conditions, the absence of a letter can be seen as evidence against the *Same* response. After all, S_2 does not perfectly match S_1. In summary, even though we did not anticipate this trend in the data, we still wanted to report our post hoc exploration for sake of transparency.

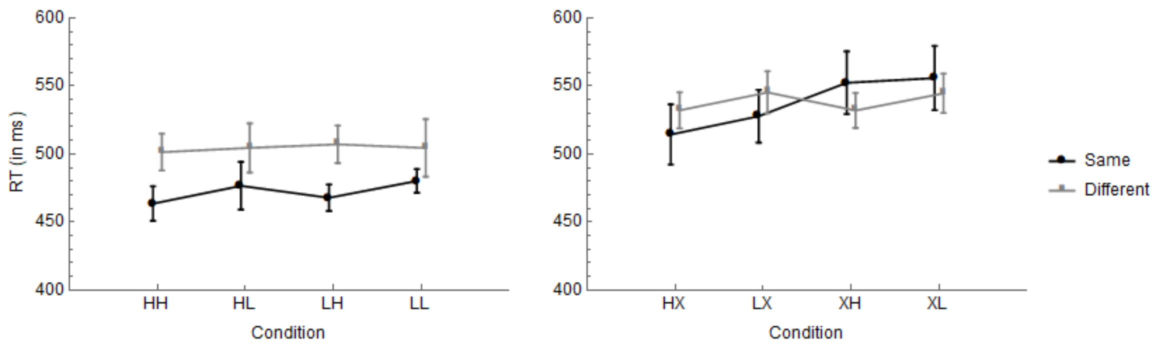


Figure 5.2. Mean response times (in ms) as a function of the experimental condition and the response (Same or Different) for correct trials. For the conditions displayed in the left plot, two letters are presented at S2, whereas for the conditions displayed in the right plot, one letter is presented at S2. Error bars are difference- and correlation-adjusted 95% confidence intervals of the mean.

Before continuing our main analyses, we also test the effectiveness of our manipulations. To do so, we used Kolmogorov-Smirnov tests for distribution ordering (Houpt et al., 2014). Such tests ensure that the distribution of the HH condition is faster than the HL and the LH conditions, and that the distribution of the LL condition is slower than the HL and the LH conditions. We conducted these tests for *Same* and *Different* trials separately. Unfortunately, in both *Same* and *Different* trials, the Kolmogorov-Smirnov tests show that the group distributions of RTs are similar between the HH and the HL / LH conditions, and between the LL and the HL / LH conditions. In other words, our manipulations on the clarity of the letters did not slow down participants as much as anticipated.

Because the manipulation was not as effective as anticipated, we cannot draw strong theoretical conclusions about the cognitive architecture underlying *Same-Different* judgments. For sake of transparency, and because we planned to conduct these analyses, we nevertheless report the MIC and SIC(t) of the group data, but we will put more emphasis on individual participant analyses, particularly on participants for which the manipulation was effective.

Cognitive Architecture

The data used in the group analysis are the mean RTs and average distributions of 19 participants in each condition. Using these measures, we conducted two SFT analyses: 1) the mean interaction contrast and 2) the survival interaction contrast. These two analyses are used to diagnose the cognitive architecture and stopping rule. Each analysis is conducted separately for *Same* and *Different* trials.

The MIC is used to estimate whether the cognitive architecture is serial, parallel or coactive. It corresponds to the difference between the two extreme conditions (HH and LL) and the two middle conditions (HL and LH). The means are illustrated in Figure 5.3. Using a 2 × 2 analysis of variance (ANOVA), we can test whether this interaction differs or not from zero, which if it does, would signify that the architecture is either parallel or coactive. For *Same* trials, the MIC = -0.70 ms, suggesting that the architecture is likely serial, $F(1, 72) < 0.01$, $p = .99$. For *Different* trials, the MIC = -5.82 ms, suggesting again that the architecture is likely serial, $F(1, 72) = 0.01$, $p = .91$.

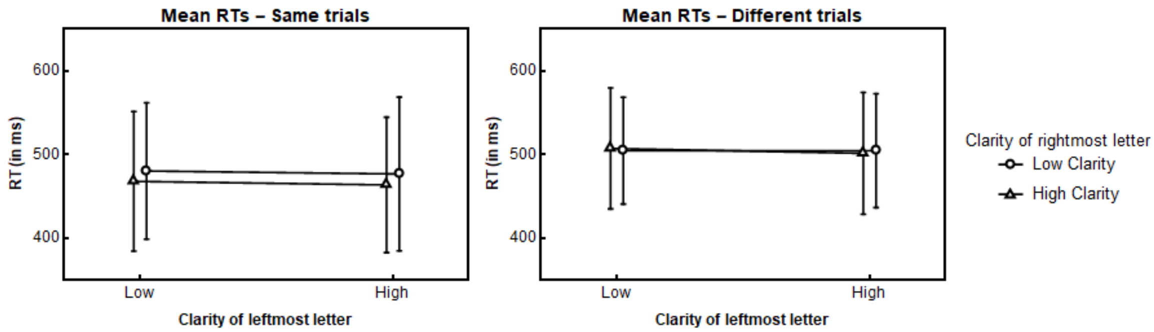


Figure 5.3. Mean response times (in ms) as a function of the clarity of the leftmost letter and the clarity of the rightmost letter, for Same trials (left plot) and Different trials (right plot) when both letters were presented. These plots illustrate the Mean Interaction Contrasts (MIC). Error bars are 95% confidence intervals of the mean.

The SIC is used like the MIC to identify which architecture characterizes the cognitive process. It further provides information about the stopping rule of the decision process, that is, either self-terming or exhaustive. The SIC curves are computed the same way as the MIC but using the distribution of RTs instead of only the mean RTs. Two Kolmogorov-Smirnov tests are used to determine whether the maximum and the minimum values of the SIC are different from zero. The group SIC curve is displayed in Figure 5.4 (in black), along with the SIC curves of individual participants (in grey or in colour). Because the average distributions are calculated using the geometric mean (Cousineau et al., 2016), the 95% confidence boundaries (in dark grey) are based on the standard errors of the geometric mean (Harding et al., 2014).

For *Same* trials, both the maximum and the minimum values of the SIC are not different from zero ($D^+ = 0.03$, $p = .912$; $D^- = -0.01$, $p = .987$). This signifies that the architecture is likely serial and that the stopping rule is likely self-terminating. In other words, participants process letters in sequence and stop their processing as soon as one

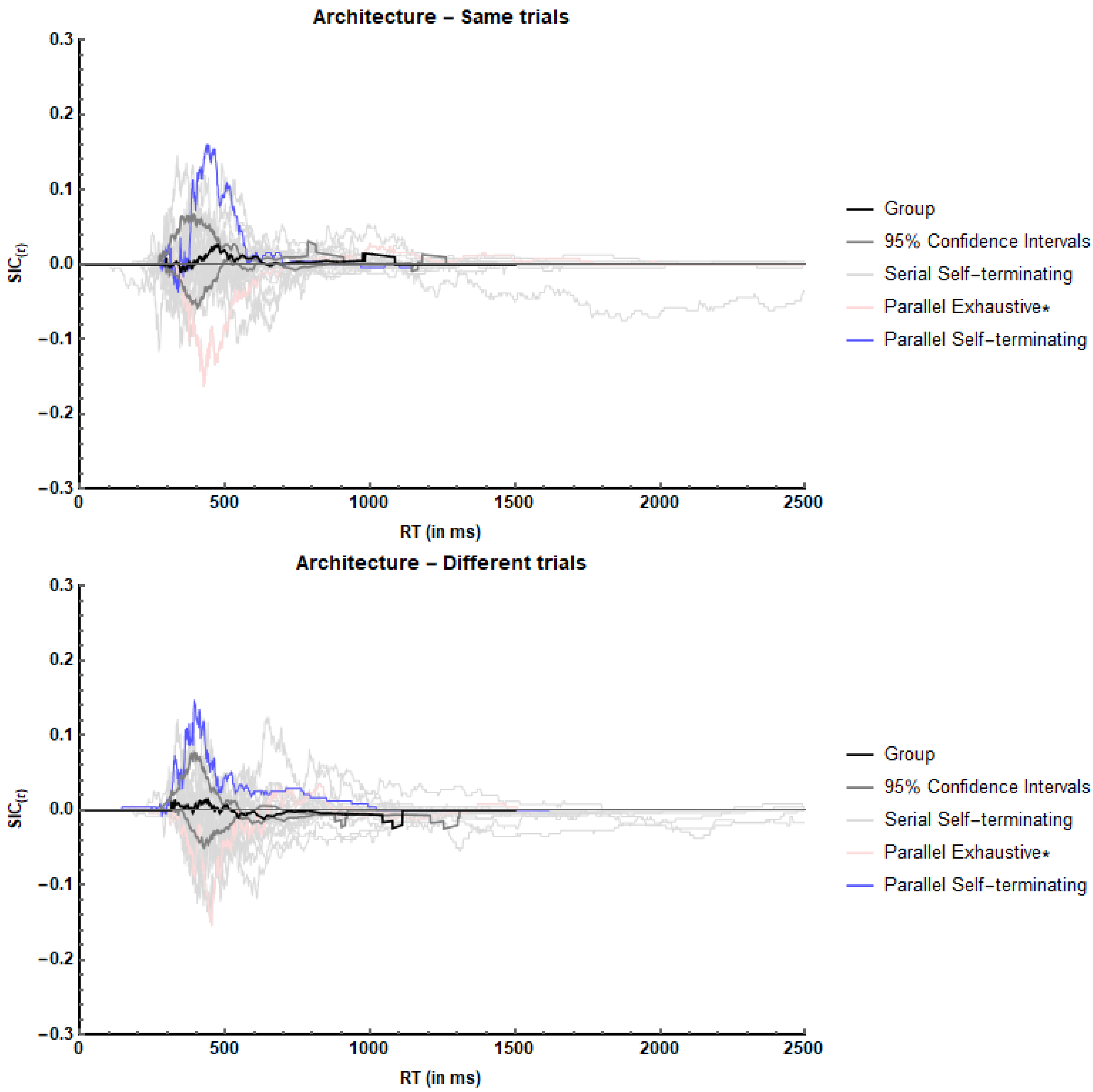


Figure 5.4. Survival Interaction Constrats (SIC(t)) for Same (top) and Different (bottom) trials. The group average is represented by the black line, accompanied by the 95% confidence intervals of the geometric mean (in dark grey). Individual participants are depicted in light grey (serial self-terminating), in light red (participant for which the SIC(t) suggest parallel exhaustive treatment, but for which the MIC suggests a serial treatment) and blue (parallel self-terminating).

letter is identified as matching. Similar results observed for *Different* trials ($D^+ = 0.02$, $p = .971$; $D^- = -0.02$, $p = .928$) suggests that processing stops as soon as one letter is identified as mismatching.

Because the blurring manipulation was not as effective as anticipated, we must consider the possibility that it precluded the detection of a parallel or coactive architecture. That said, for some participants, the manipulation was effective. For instance, in the *Same* condition, Kolmogorov-Smirnov tests show that blurring the letter slowed down participants 5, 7, 8, 9 and 18. Participant 8 exhibited a parallel self-terminating architecture in this condition. The SIC(t) curve of this participant is highlighted with the blue colour, in the top panel of Figure 5.4. The other four participants exhibited a serial self-terminating architecture. We also highlight the results of another participant (15), for who even though the MIC suggests a serial architecture, the Kolmogorov-Smirnov tests of their SIC(t) curve show that this participant might have exhibited a parallel exhaustive treatment of the stimuli. This participant is highlighted in the top panel of Figure 5.4 with a light red colour.

In the *Different* condition, the blurring manipulation did not slow down any participant. That said, we were still able to detect that participant 18 used a parallel self-terminating architecture. Also, like was the case for participant 15 in the *Same* condition, participant 13 in the *Different* condition, despite a MIC suggesting a serial architecture, exhibited a parallel exhaustive treatment when looking at their SIC(t) curve. This participant is highlighted with the light red colour in the bottom panel of Figure 5.4.

Table 2 provides a summary of individual SFT tests for every participant. Again, since the manipulation was not as effective as anticipated, we cannot draw strong

Table 5.2. Summary of individual participant analyses.

Subject	Condition	MIC	SIC(t)+	SIC(t)-	Capacity	Summary
1	Same	2.87, $p = .918$	0.03, $p = .871$	0.10, $p = .276$	-3.06, $p = .002$	Serial Self-Terminating; Limited Capacity
	Different	7.09, $p = .865$	0.06, $p = .600$	0.13, $p = .138$	-7.98, $p < .001$	Serial Self-Terminating; Limited Capacity
2	Same	7.54, $p = .593$	0.05, $p = .771$	0.11, $p = .215$	0.30, $p = .767$	Serial Self-Terminating; Unlimited Capacity
	Different	-9.14, $p = .496$	0.02, $p = .968$	0.09, $p = .367$	-4.61, $p < .001$	Serial Self-Terminating; Limited Capacity
3	Same	23.46, $p = .168$	0.10, $p = .274$	0.06, $p = .597$	-2.71, $p = .007$	Serial Self-Terminating; Limited Capacity
	Different	-15.79, $p = .343$	0.04, $p = .792$	0.07, $p = .495$	-7.09, $p < .001$	Serial Self-Terminating; Limited Capacity
4	Same	-3.22, $p = .846$	0.04, $p = .789$	0.07, $p = .524$	-2.67, $p = .008$	Serial Self-Terminating; Limited Capacity
	Different	-7.04, $p = .699$	.037, $p = .849$	0.07, $p = .575$	-8.20, $p < .001$	Serial Self-Terminating; Limited Capacity
5	Same	15.77, $p = .373$	0.06, $p = .624$	0.06, $p = .683$	0.21, $p = .836$	Serial Self-Terminating; Unlimited Capacity
	Different	6.85, $p = .709$	0.06, $p = .656$	0.03, $p = .903$	-8.50, $p < .001$	Serial Self-Terminating; Limited Capacity
6	Same	-14.68, $p = .594$	0.02, $p = .968$	0.07, $p = .545$	-1.85, $p = .065$	Serial Self-Terminating; Unlimited Capacity
	Different	25.74, $p = .210$	0.13, $p = .152$	0.01, $p = .152$	-7.28, $p < .001$	Serial Self-Terminating; Limited Capacity
7	Same	-0.66, $p = .957$	0.06, $p = .618$	0.08, $p = .441$	3.30, $p < .001$	Serial Self-Terminating; Super Capacity
	Different	-38.51, $p = .101$	0.02, $p = .952$	0.14, $p = .078$	-5.25, $p < .001$	Serial Self-Terminating; Limited Capacity
8	Same	24.59, $p = .041$	0.16, $p = .035$	0.04, $p = .843$	-1.26, $p = .208$	Parallel Self-Terminating; Unlimited Capacity
	Different	2.30, $p = .821$	0.06, $p = .682$	0.08, $p = .428$	-4.02, $p < .001$	Serial Selft-Terminating; Limited Capacity
9	Same	-3.81, $p = .853$	0.11, $p = .204$	0.04, $p = .822$	-4.46, $p < .001$	Serial Self-Terminating; Limited Capacity
	Different	-23.64, $p = .435$	0.01, $p = .979$	0.10, $p = .302$	-5.35, $p < .001$	Serial Self-Terminating; Limited Capacity
10	Same	3.37, $p = .733$	0.11, $p = .196$	0.05, $p = .723$	-3.41, $p < .001$	Serial Self-Terminating; Limited Capacity
	Different	-10.79, $p = .333$	0.03, $p = .877$	0.07, $p = .526$	-4.45, $p < .001$	Serial Self-Terminating; Limited Capacity

(Table continues next page)

11	Same	-94.77, *p* = .201	0.04, *p* = .865	0.09, *p* = .389	-5.09, *p* < .001	Serial Self-Terminating; Limited Capacity
	Different	-20.14, *p* = .769	0.12, *p* = .162	0.06, *p* = .662	-5.81, *p* < .001	Serial Self-Terminating; Limited Capacity
12	Same	-1.79, *p* = .849	0.07, *p* = .552	0.12, *p* = .211	-1.69, *p* = .090	Serial Self-Terminating; Unlimited Capacity
	Different	2.62, *p* = .831	0.12, *p* = .239	0.04, *p* = .875	-8.18, *p* < .001	Serial Self-Terminating; Limited Capacity
13	Same	-14.20, *p* = .419	0.02, *p* = .948	0.08, *p* = .263	-5.11, *p* < .001	Serial Self-Terminating; Limited Capacity
	Different	-16.59, *p* = .408	0.03, *p* = .882	0.16, *p* = .048	-8.39, *p* < .001	Serial Self-Terminating; Limited Capacity
14	Same	-0.10, *p* = .994	0.07, *p* = .520	0.06, *p* = .668	-0.93, *p* = .351	Serial Self-Terminating; Unlimited Capacity
	Different	5.04, *p* = .735	0.10, *p* = .308	0.04, *p* = .818	-5.83, *p* < .001	Serial Self-Terminating; Limited Capacity
15	Same	-23.38, *p* = .357	0.02, *p* = .929	0.17, *p* = .032	-3.24, *p* = .001	Serial Self-Terminating; Limited Capacity
	Different	-3.75, *p* = .861	0.07, *p* = .507	0.07, *p* = .498	-6.51, *p* < .001	Serial Self-Terminating; Limited Capacity
16	Same	21.93, *p* = .554	0.06, *p* = .678	0.10, *p* = .228	-5.71, *p* < .001	Serial Self-Terminating; Limited Capacity
	Different	-8.67, *p* = .826	0.05, *p* = .767	0.04, *p* = .801	-7.45, *p* < .001	Serial Self-Terminating; Limited Capacity
17	Same	9.42, *p* = .506	0.11, *p* = .258	0.11, *p* = .244	-4.26, *p* < .001	Serial Self-Terminating; Limited Capacity
	Different	-2.99, *p* = .851	0.04, *p* = .834	0.12, *p* = .150	-6.51, *p* < .001	Serial Self-Terminating; Limited Capacity
18	Same	14.43, *p* = .153	0.15, *p* = .070	0.04, *p* = .815	3.09, *p* = .002	Serial Self-Terminating; Super Capacity
	Different	29.92, *p* = .005	0.15, *p* = .066	0.01, *p* = .991	-9.08, *p* = .001	Parallel Self-Terminating; Limited Capacity
19	Same	6.40, *p* = .747	0.07, *p* = .580	0.04, *p* = .779	3.31, *p* = .001	Serial Self-Terminating; Limited Capacity
	Different	-9.03, *p* = .559	0.03, *p* = .877	0.08, *p* = .490	-5.80, *p* < .001	Serial Self-Terminating; Limited Capacity
20	Same	11.77, *p* = .393	0.13, *p* = .121	0.02, *p* = .946	-5.98, *p* < .001	Serial Self-Terminating; Limited Capacity
	Different	-21.42, *p* = .125	0.02, *p* = .969	0.12, *p* = .168	-5.02, *p* < .001	Serial Self-Terminating; Limited Capacity

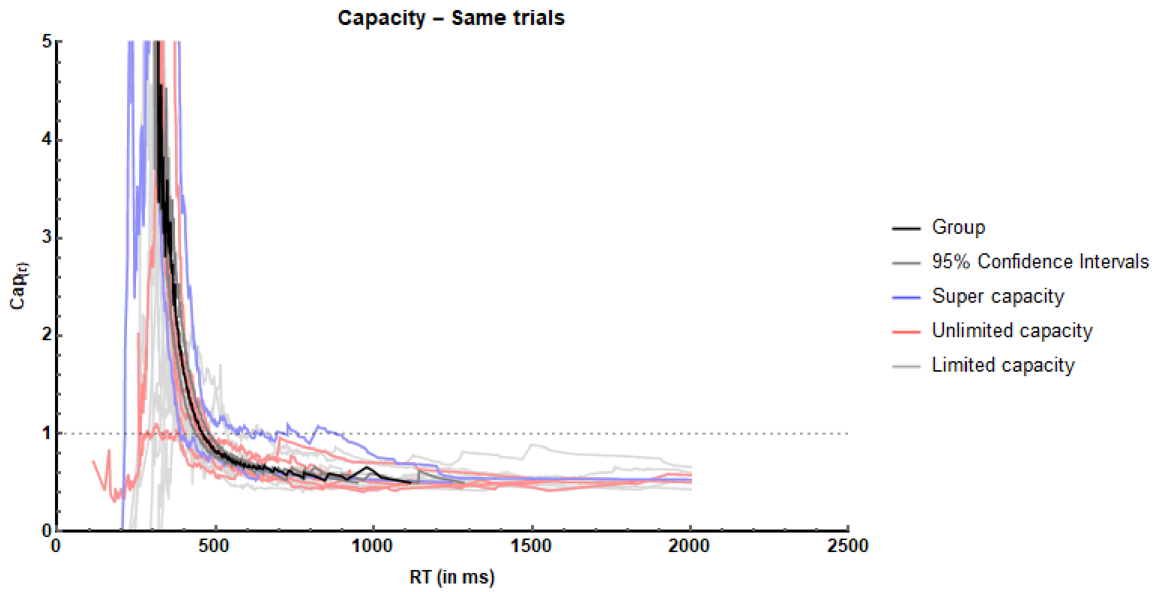
Notes. The column MIC reports the mean interaction contrast and the p-value of the interaction. The column SIC(t)+ reports the Kolmogorov-Smirnov D statistic and the p-value of the test for the positive portion of the curve and the column SIC(t)- for the negative portion of the curve. The column Capacity reports the capacity coefficient and the p-value of the Kolmogorov-Smirnov test.

conclusions regarding the cognitive architecture underlying *Same-Different* judgments. That being said, we can focus on participants for which the manipulation was effective to gather at least some direction. Despite their differences, the aforementioned participants do not seem to have faster RTs or better accuracy compared to the other participants. This shows that not all participants approach the task the same way, but that the most predominant architecture was serial, that is processing one letter at a time, for both the *Same* and the *Different* trials. In light of these results, we can argue that the cognitive architecture does not seem to explain why *Same* trials are faster than *Different* trials.

Capacity Coefficient

The capacity coefficient is used to estimate whether the efficiency of the cognitive system is limited, unlimited or super capacity. It is computed by comparing the distribution of RT in the HH condition to an estimated distribution for an unlimited capacity using the HX and the XH conditions. Using a z test, we can assess the observed effect of increased workload compared to this predicted null effect (unlimited capacity). The group results are summarized in Figure 5.4 (in black) along with individual capacity coefficient curves (in grey and colour). The dashed line represents theoretical unlimited capacity. The 95% confidence intervals (in dark grey) are calculated again using the standard error of the geometric mean (Harding et al., 2014).

For *Same* trials, although the capacity tends to be limited, the statistical test shows that the overall capacity is unlimited (z = -1.37, p = .172). Looking at Figure 5.5 to understand this discrepancy, we see that the capacity coefficient does not reach an asymptote until about 500 ms; before this time, it seems strongly super-capacity. This suggests that participants might benefit from redundancy at first but are quickly hindered by the presence of additional letters. In contrast, for *Different* trials, the capacity is immediately limited, as supported by the statistical test ($z = -3.86$, $p < .001$). When the letters are different, participants are systematically hindered by the increase in workload.



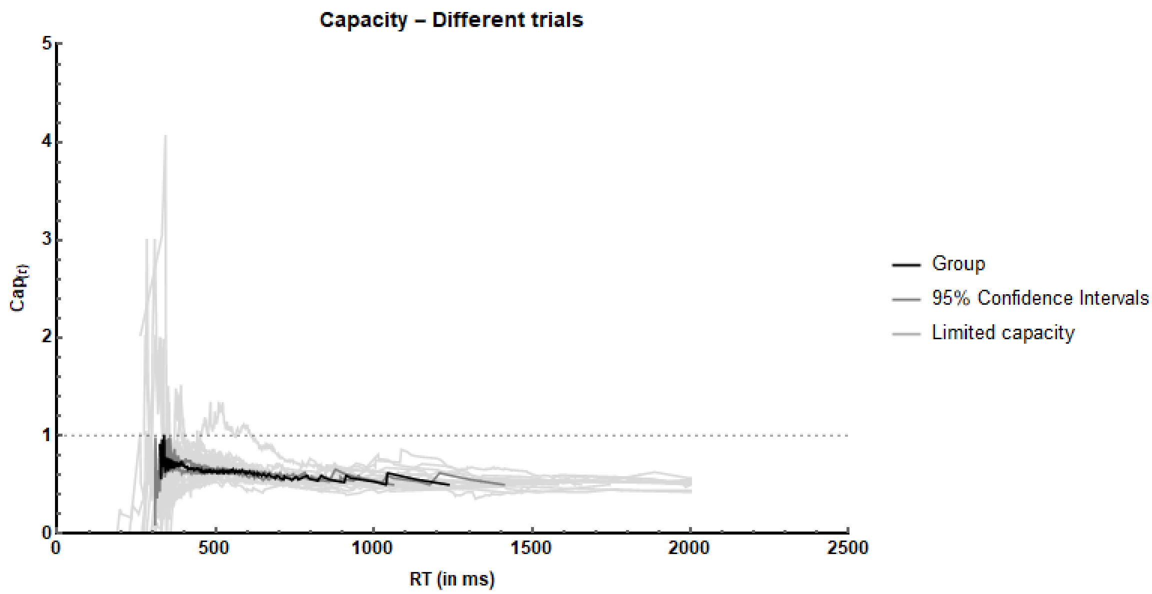


Figure 5.5. Survival Interaction Constrats (SIC(t)) for Same (top) and Different (bottom) trials. The group average is represented by the black line, accompanied by the 95% confidence intervals of the geometric mean (in dark grey). Note that the confidence intervals are close to the group curve and might not be obvious at first glance. Individual participants are depicted in light grey (limited capacity), in red (unlimited capacity) and blue (super-capacity).

It is typical to observe large oscillations in the capacity coefficient in the first few hundred milliseconds. However, looking at Figure 5.5, it is clear that the differences between the *Same* capacity and the *Different* capacity curves are not caused by a statistical fluke. For instance, the group capacity coefficient curve for *Same* trials intersects with the unlimited capacity reference line (Cap(t) = 1) at 457 ms, whereas the group capacity coefficient curve for *Different* trials is always below the unlimited capacity reference line. This discrepancy in capacity coefficients between *Same* and *Different* trials could potentially explain why *Same* responses are faster than *Different* responses. Looking at individual participants, whereas they all showed a limited capacity in *Different* trials, some participants had unlimited or super-capacity in the *Same* trials (listed in Table 5.2).

Let us return to the unexpected result in trials with only one letter presented at S_2. Recall that in *Same* trials, participants were faster to respond when the leftmost letter was presented compared to when the rightmost letter was presented, an effect not observed in *Different* trials. We argued earlier that there might be more response competition in these trials, which inhibits the *Same* response. A consequence of this response competition is that it would decrease the capacity advantage that benefits *Same* responses. This would explain why the capacity coefficient decreases to reach an asymptote at limited capacity as it does in the *Different* trials.

Discussion

The two main objectives of this study were to formally assess the cognitive architecture and processing capacity of the system underlying *Same-Different* judgments. These objectives stem from two roots. First, one result of the *Same-Different* task that has been replicated many times is the fast-same effect, whereby participants respond faster in the *Same* condition compared to the *Different* condition. This discrepancy might be explained by different cognitive architectures for the systems underlying *Same* and *Different* judgments. For example, Bamber (1969) proposed that *Same* responses are faster because the underlying system works in parallel, whereas the underlying system of *Different* responses works in serial. The fast-same effect might also be caused by the capacity of the systems, that is how efficiently it can process the information as a function of the workload (the amount of information that needs to be treated). Expanding on Bamber (1969), Decker (1974) argued that not only *Different* trials were processed serially, but that they were also subject to a limited capacity. As discussed in the introduction, researchers assumed different architectures and capacity when explaining the fast-same effect (also see Table 5.1). This leads to the second question: Which theory assumed architecture and capacity correctly?

The results of the current study are – unfortunately – not conclusive enough to make any strong claim regarding cognitive architecture, because the blurring manipulation was not as effective as anticipated. At best, the results suggest that in both *Same* and *Different* trials, participants processed the letters serially, but that some participants processed the letters in parallel. The serial architecture was originally assumed by many researchers (i.e., Krueger, 1978; Nickerson, 1965). Bamber (1969) and Decker (1974) were also arguing for a serial treatment, but only for *Different* trials (they argued that participants processed *Same* trials in parallel, or *holistically*). Hawkins (1969) highlighted that there was no need for a serial treatment, and that a parallel model could also account for the fast-same effect. One conclusion that we can draw from the results of this study is that the fast-same effect is not dependant on the underlying cognitive architecture: Different participants had different architecture without apparent effect on the presence of the fast-same effect.

Another critical assumption made by most researchers is that the workload does not affect the processing efficiency of the system underlying *Same-Different* judgments. In our study, we observed that most participants had an overall limited capacity in *Different* trial. Only a few theories accounted for this hindrance in capacity for the *Different* trials (Decker, 1974; Hawkins & Shigley, 1972; Link & Tindall, 1971). That said, these researchers also argued that there was some form of parallel processing in the *Same* condition. As for *Same* trials, i) the capacity coefficient asymptote is reached less rapidly, and ii) there exists greater individual differences. The results show that in the first 500 ms, participants are more efficient at processing *Same* information compared to *Different* information. However, possibly due to response competition, participants capacity recedes towards limited capacity. This initial, short lasting, but strong boost in processing capacity could explain why *Same* responses are faster than *Different* responses. This is novel result, as no theory ever explicitly attributed the fast-same effect solely to processing capacity.

Capacity and the Fast-Same Effect

Miller's model (1978, 1982) stands out as an outlier in the *Same-Different* task models, proposing that the system underlying the task might work in coactivity and in super-capacity. Whereas the current study does not support this theory, we observe an interesting trend in the capacity coefficients between 300 and 500 ms, which was seen in Figure 5.5. Indeed, for *Different* trials, the capacity coefficient is fairly flat and reaches an asymptote at approximately 400 ms). For *Same* trials however, we see that the capacity coefficient is initially super-capacity, then decreasing between 300 and 500 ms, after which it reaches its asymptote. This signifies that between 300 and 500 ms, the processing capacity is larger in *Same* trials than in *Different* trials. To our knowledge, this *time-limited capacity enhancement* in *Same* trial was never proposed as a potential explanation for the fast-same effect. This is probably because processing capacity was often ignored in models of the *Same-Different* task. That said, we can draw some parallels with Farell (1985)'s model, which highlights the role of attention in the fast-same effect.

In neuropsychological studies, one event-related potential that occurs at around 300 ms after the presentation of a stimulus is the P300. This event-related potential has been associated with workload in short-term memory (Picton, 1992). Typically, a large amplitude of the P300 is associated with a more important workload (i.e., in n-back tasks, Scharinger, Soutschek, Schubert, & Gerjets, 2017). It could also be associated with participants paying more attention to an unexpected stimulus (Picton, 1992). This resonates with Farell (1985)'s model of the *Same-Different* task. Farell proposed that a single process is underlying the task, but that it is modulated by attention (in the sense that *Different* trials require more attention than *Same* trials because participants are expecting matching stimuli). Perhaps this is mirrored by the lack of processing capacity in the early stages of the processing (the first 500 ms) in *Different* trials, as observed in this study.

It could also explain the unexpected result obtained in trials in which only one letter was presented at S_2 (faster mean RTs when the leftmost letter was presented for

Same but not for *Different* trials. If participants rely on pre-attentive processing for *Same* trials, they might resort to their default serial scanning. On *Different* trials however, participants already spend attentional resources to process the stimulus and might be better at directing their processing resources at a specific area on the screen.

The potential role of the P300 in the *Same-Different* task remains unclear. One could expect that a larger amplitude of this event-related potential would be associated with more processing capacity. However, it could also be associated with a need for additional attentional resources, which would consequently reduce processing capacity. Indeed, some researchers showed that stimulus treatment and attention are likely in competition for the same resources (Swan & Wyble, 2014; Wyble et al., 2011). If *Different* responses require more attention than *Same* responses, as suggested by Farell, this could explain why we observe that *Same* responses benefits, initially, from more processing resources, therefore, from more processing capacity.

Limitations of the Current Study

One of the main limitations of the study is that the manipulation (i.e., blurring the letters), was not as effective as anticipated. It is difficult to interpret why this occurred. In principle, participants should be slower to encode/process blurry information compared to clear information. This was not reflected in the Kolmogorov-Smirnov tests conducted to test the efficiency of this manipulation. Perhaps a future replication study could be conducted with the letters being blurred even more. That said, we must consider the possibility that increasing the blurriness of the letters too much might render them undetectable from one another, which would result in massive error rates.

Another limitation is the fact that the number of conditions could be considered rather low from *Same-Different* standards. Indeed, it is not rare to see researchers vary the number of elements (in our case letters) from one to four and the number of mismatching elements from zero to four. For example, Bamber (1969) used fourteen conditions in total (including intermediate conditions such as three letters, one mismatching letter) , whereas we used – excluding the conditions in which blurred letters were used – four conditions (one letter, one match or one mismatch; two letters, two matches or two mismatches). We only included these conditions because in addition to our other manipulations, this represented four sessions of one hour for every participant. However, we strongly encourage researchers to include intermediate conditions such as two letters, one mismatching letter (and one matching letter) in future designs.

Final Remarks and Future Directions

Research on the *Same-Different* task blossomed in 1960s and the 1970s, but slowed down in the 1980s, probably due to the plethora of proposed models and the lack of consensus. We identified two unsupported assumptions that present a point of discord between models: cognitive architecture and processing capacity. In order to be as objective as possible in answering this question, we preregistered an experimental protocol that aimed to directly assess those two research questions. We found that the cognitive architecture does not seem to influence the fast-same effect.

We also found that *Different* trials were characterized by limited processing capacity, whereas most models of the *Same-Different* task assumed (by default) that the processing capacity was unlimited (the few exceptions being Decker, 1974; Hawkins & Shigley, 1972 and Link & Tindall, 1971). In contrast, *Same* trials were characterized by a more ambiguous processing capacity, as illustrated in Figure 5.5. Overall, the system was working somewhere between unlimited and limited capacity. However, between 300 and 500 ms, we found that the capacity coefficient in *Same* trials is super-capacity, after which it reaches an asymptote.

Future work should focus on this discrepancy in processing capacity between *Same* and *Different* trials. A good approach would be to use electroencephalogram (EEG) data and explore event-related potentials around 300-500 ms (such as the P300) in a *Same-Different* design. Specifically, it would be pertinent to see if the P300 amplitude varies between *Same* and *Different* trials.

Appendix

Mean Interaction Contrast (MIC)

Using the double-factorial design, mean RTs can be computed for the HH, HL, LH and LL conditions. Using these four means, the Mean Interaction Contrast (MIC) can be calculated using this formula:

$$MIC = (\overline{RT}_{HH} + \overline{RT}_{LL}) - (\overline{RT}_{HL} + \overline{RT}_{LH}) \quad (1)$$

The MIC can provide insightful information about the architecture of the cognitive system. If the system is working in serial, the slowdown in processing caused by the L manipulation on isolated attributes (HL and LH conditions) should stack in the LL condition. Therefore, the MIC should be equal or close to 0. If the MIC is lower than 0, it suggests that the architecture is parallel, and the stopping rule is exhaustive. Indeed, the probability of the two channels completing their processing after *t* ms is decreased when an attribute is in the L condition. The probability of one channel being very slow increases when the two dimensions are in the L condition. Conversely, the probability of either of the two channels completing their processing after *t* ms is increased when an attribute is in the H condition. The probability of a channel being very fast increases when the two attributes are in the H condition. Therefore, when the architecture is parallel, and the stopping rule is self-terminating, the MIC is greater than 0. When the architecture is coactive, the MIC is also greater than 0, because an increase in the channel processing rates signifies that the common threshold will be reached faster.

Despite being a useful tool, the MIC is limited to the mean RT and cannot distinguish between stopping rules for serial architecture, or between a parallel self-terminating and a coactive system (Harding et al., 2016). To alleviate these limitations, the SFT uses the whole distribution of RTs.

Survival Interaction Contrast (SIC)

The probability density function (PDF) corresponds to the density of observing a response at precisely *t* ms. Another way of estimating the distribution of RT is to compute the probability of observing a response before *t* ms, that is the cumulative density function (CDF). Alternatively, one could compute the probability of no response being observed before *t* ms, that is 1 – CDF, called the survivor function (SF). The SF is a useful tool to illustrate a mental process and can be portrayed as the probability of the cognitive system still being active (or still accumulating information) after *t* ms.

Analogous to the computation of the MIC, the *Survivor Interaction Contrast* (SIC) can be calculated for every time point of *t* (in ms). Using the SF of all four conditions (HH, HL, LH and LL), the SIC(*t*) can be obtained using this formula:

$$SIC(t) = \left(SF_{HH}(t) + SF_{LL}(t)\right) - \left(SF_{HL}(t) + SF_{LH}(t)\right) \qquad (2)$$

Plotting the SIC(t) curve provides the researcher with a diagnostic tool to detect the architecture and the stopping rule of a cognitive system (Townsend & Nozawa, 1995). Figure 5.A1 pictures the five distinct SIC(*t*) signature curves. A Kolmogorov-Smirnov test can be used to evaluate if the minimum and the maximum are different from zero (Houpt & Burns, 2017; Houpt & Townsend, 2010).

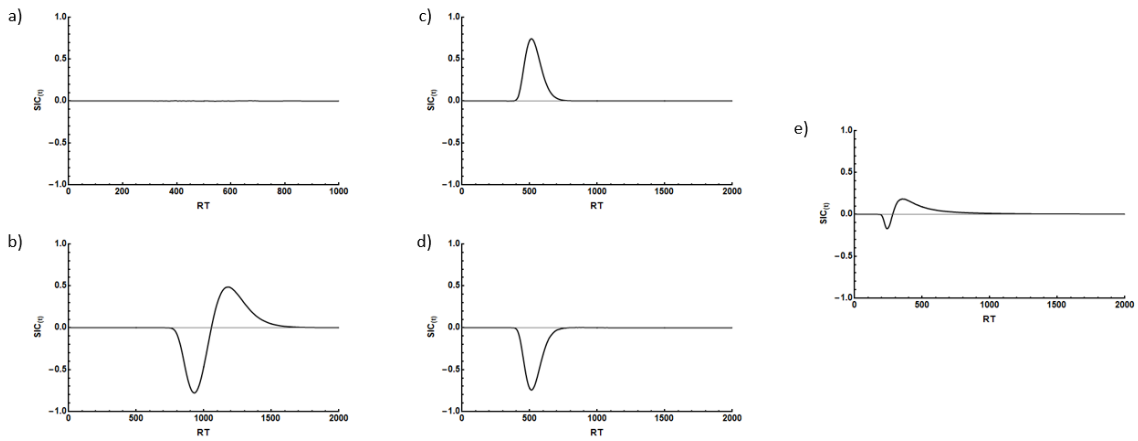


Figure 5.A1. Survival interaction contrasts (SIC(t)) signature curves for all five architectures and stopping rule combinations. The panel a) corresponds to a system working in serial with a self-terminating stopping rule. The curve resembles a flat line and does not deviate from 0. The panel b) corresponds to a system working in serial with an exhaustive stopping rule. The negative and the positive areas under the curve differ from zero but are equal. The panel c) corresponds to a system working in parallel with a self-terminating stopping rule. Only the positive portion is different from zero. The panel d) corresponds to a system working in parallel with an exhaustive stopping rule. Only the negative portion is different from zero. Finally, the panel e) corresponds to a system working in coactivity. The curve resembles the serial exhaustive curve, as both the positive and the negative portions are different from zero. However, the area under the curve in the positive portion is higher than the area under the curve in the negative portion.

A serial self-terminating architecture is characterized by a flat SIC(t) curve, where neither the positive nor the negative portion are different from zero. A serial exhaustive architecture is characterized by a negative portion at the beginning of the curve and a positive portion at the end of the curve. A parallel self-terminating architecture is characterized by a negative portion only, whereas a parallel exhaustive architecture is characterized by a positive portion only. Finally, the coactive architecture show similarity with the serial exhaustive curve, starting with a small negative portion, then with a larger positive portion. The MIC should be used to distinguish between these two architectures.

Although the SIC(*t*) signature curves are shown to be a powerful tool to detect architecture and stopping rules, some situations might lead to misdiagnosis. If the decision thresholds of a system are highly variable, SIC(*t*) curves might mimic a parallel self-terminating architecture, as shown by Harding, LeBlanc, Goulet, & Cousineau (2017). Misdiagnosis might also arise from noisy data caused by non-equivalent manipulations on the two attributes. Indeed, if the salience of each source of information is inappropriately balanced in the factorial manipulation, one attribute could be overly represented in some conditions and add noise to the SIC(*t*) curves (Houpt & Fifić, 2017). As SIC(*t*) curves are estimated from a finite number of trials, the estimated SF curves can be sensible to asymmetric manipulation design. Finally, some researches have recently suggested that participants might use mixtures of architecture throughout the experimental session. In other words, some participants might use a parallel strategy, say for 60% of the trials, and a serial strategy for the remaining 40% of the trials (Little, Nosofsky, & Denton, 2011; Moneer, Wang, & Little, 2016). Although SIC(*t*) curves can still be used to test the plausibility of mixture architectures, they must be compared with the predictions made by mixture models fitting.

Capacity Coefficient

To answer questions regarding the capacity of the cognitive system, SFT uses conditions where only one attribute is presented to the participants. Essentially, capacity assess whether increasing the workload of the task (increasing the number of active channels or the complexity of the stimulus) affects the processing rates of the channel.

SFT measures the capacity coefficient by comparing conditions where one attribute is present to conditions where two attributes are present.

The computation of the capacity coefficient requires the use of cumulative hazard functions (CHF), which correspond to the probability at *t* ms of the response occurring on the next iteration, given that no response has been given before *t* ms. Alternatively, it can be described as the probability that a state change is imminent (Chechile, 2011).

The CHF of three conditions must be estimated to calculate the capacity coefficient in self-terminating designs (OR designs): two conditions in which the single dimensions are presented (HX and XH) and the redundant condition (HH) (Altieri et al., 2017; Harding et al., 2016; Houpt et al., 2014). To compute the capacity coefficient of an OR design, the following formula must be used:

$$C_{OR} = \frac{CHF_{HH}(t)}{CHF_{HX}(t)+CHF_{XH}(t)} \qquad (3)$$

In cases of exhaustive designs (AND designs), where both dimensions must be obligatory assessed before providing a response, the capacity coefficient is computed using reversed hazard functions (K) instead of CHF (Altieri et al., 2017; Harding et al., 2016; Houpt et al., 2014). To compute the capacity coefficient of an AND design, the following formula must be used:

$$C_{AND} = \frac{K_{HX}(t)+K_{XH}(t)}{K_{HH}(t)} \quad (4)$$

For the sake of simplicity, we will further refer to the adequate capacity coefficient – C_{OR} or C_{AND} – as C(t). If the value of C(t) is equal to 1, the system works with full capacity. Adding more attributes does not increase nor decrease the capacity of the channels compared to conditions where only one attribute is shown. If the value of C(t) is lower than 1, the system works with limited capacity. The increase in dimensions hinders the processing rate of the channels compared to conditions where only one dimension is shown. Finally, if the value of C(t) is greater than 1, the system works with super capacity. Herein, more attributes speed up the processing rate of the channels in comparison to conditions where only one attribute is presented. Figure 5.A2 illustrates capacity coefficient signatures of full, limited and super capacity systems.

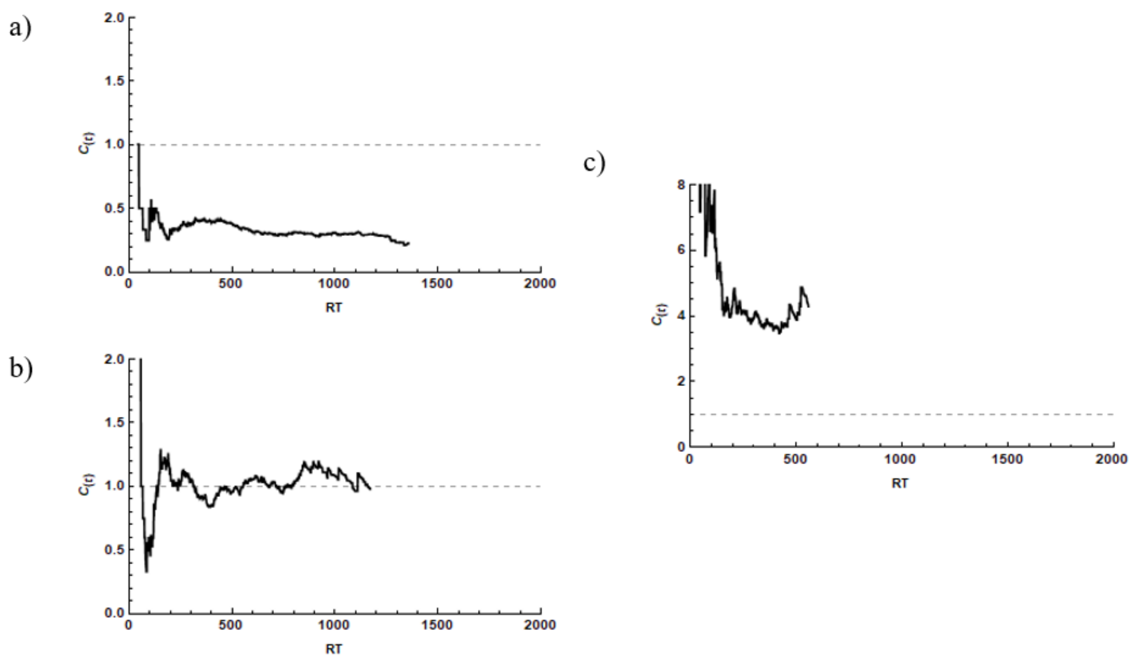


Figure 5.A2. Capacity coefficient (C(t)) signature curves for all three workload capacity types. Each plot contains a reference line at C(t) = 1, corresponding to the system working at full capacity. The panel a) corresponds to a system working at limited capacity. The curve is below the reference line. The panel b) corresponds to a system working at unlimited capacity. The curve is directly on the reference line except for oscillations for small RTs owing to near-zero counts. The panel c) corresponds to a system working at super capacity. The curve is above the reference line. When the system benefits a lot from additional channels, C(t) will likely have very high values (> 1000).

The capacity coefficient can be statistically tested using a *z* test on any given value of *t* (Houpt & Burns, 2017). Although the test can be repeated for every value of *t*, it is recommended to limit the number of tests to avoid the increased risk of committing a Type I error. Thus, the capacity coefficient should be tested using a log-rank test, as described in Houpt & Townsend (2012).

Chapter 6 – Judgment Day

One hundred thirty-seven thousand four hundred forty response times later, we now have a better understanding of how humans do *Same-Different* judgments. In this chapter, I summarize the main findings of the studies reported in this thesis and critically assess how they support or falsify the various models presented in the first chapter. Finally, at the light of these new results, I propose future directions for researchers that wish to continue studying the *Same-Different* task and the fast-same effect.

What is the fast-same effect?

This broad question is behind every study on *Same-Different* judgments, not just the ones presented in this thesis. However, since there are multiple plausible answers, we need to address the question from multiple angles. The first angle is to determine if the effect is induced by the participant or by the stimulus. The results of this thesis highly suggest that the fast-same effect is caused by the stimulus, not by the participant. In Study 1, we observed the fast-same effect even thought participants pressed the same answer key to respond *All-Same* and *All-Different*. The fit of the LBA to the data showed that, indeed, participants were equally biased to press the *All-Same* and the *All-Different* key, which is not surprising since in both conditions, they need to press the same physical answer key. Without informing the model about this particularity in the design, the LBA was still able to capture this key-specific bias.

Instead, the LBA attributed the discrepancy between the *All-Same* response times and the *All-Different* response times to another parameter: accumulation rate. Participants were faster to process *All-Same* stimuli than any other stimuli, including *All-Different* stimuli. In the *All-Same* condition, participants already have a mental representation of the correct response, an identical replica of the stimulus presented 400 ms earlier. This likely boosted the processing speed of matching elements, resulting in faster mean response time in the *All-Same* condition. However, this processing boost seems to only occur when participants can represent the stimulus mentally, not just when the response is *Same*. For instance, in Study 2, participants only showed a fast-same effect when the letters were identical, not when they were similar. The difference between the two conditions is that in the former, participants know what an identical letter looks like, and can realize a one-to-one mapping, but in the latter, participants do not know exactly what a similar letter looks like, and must resort to a many-to-one mapping strategy.

Another aspect of Study 1 is that participants seemed to have pondered the response alternatives in a serial fashion. The distribution of response times suggests that participants followed a decision tree to realize the task. First, they tried to map the letters to their mental representation and if the mapping was successful, they responded *All*. Second, they segmented their mapping, looking if either letter was matching, and if it was the case, they responded *Some*. Finally, if neither mapping was successful, they could respond *All*, knowing that all the letters were mismatching. Even though the task in Study 1 was not a classical *Same-Different* task, it revealed that participants tend to be serial when considering multiple response alternatives.

The cognitive architecture of the *Same-Different* judgment is the second angle from which I approached the fast-same effect. Specifically, in Study 2, I addressed the hypothesis that there exists a priority for the treatment of information regarding similarity. Whereas most participants processed similarity and clarity in a serial fashion, they did so not in a fixed order. In other words, they did not process the similarity of the stimuli in priority over the clarity of the stimulus. In Study 3, cognitive architecture is once again tested, but this time at a more local level, namely how specific channels process the elements contain in the stimuli. One possible explanation for the fast-same effect was that participants process matching elements in parallel, but mismatching elements in serial. The results of Study 3 were unfortunately not strongly conclusive regarding the cognitive architecture, mainly because the experimental manipulation failed for most participants. That said, the predominantly observed architecture of the participants for which the manipulation worked was serial.

In both Study 2 and Study 3, some participants exhibited a parallel architecture, but still manifested a fast-same effect. These results do not mean that cognitive architecture is not important to understand *Same-Different* judgments. After all, the results of all three studies provide evidence in favour of a serial treatment. That said, it does not seem to be the cause of the fast-same effect.

An unexpected result observed in Study 2 is that participants used a self-terminating stopping rule despite the task requiring an exhaustive treatment of similarity and clarity. This means that participants likely only processed thoroughly one type of information to provide their response. As ineffective as this strategy seems, participants nevertheless performed the task with high accuracy. A plausible explanation is that participants first process global information contained in the stimulus and used that fast treatment to guide their more local-based treatment. This first form of treatment is likely pre-attentive, whereas the second form of treatment involves attention and is therefore more demanding on resources. Consequently, the third angle of this thesis focuses on the effect of workload on processing capacity.

In Study 3, participants quickly reached a point in which their processing capacity was highly limited due to a lack of available resources. For instance, in the *Different* condition, participants reached the *limited capacity* asymptote quite rapidly. However, in the *Same* condition, the asymptote was reached after around 500 ms of treatment. This suggests that in the first 500 ms, participants are more efficient when processing *Same* stimuli compared to *Different* stimuli. Thus, the participants use fewer attentional resources to process *Same* information and therefore, can use the saved resources to boost their processing speed.

Overall, the results of these three studies show that the fast-same effect is not a matter of participants' preference for similarity, but rather a consequence of how more resource-efficient processing *Same* information is compared to processing *Different* information.

Models Trial

At the end of his book chapter reviewing the *Same-Different* literature, Saul Sternberg mentions that "*intriguing puzzles remain to be solved*" (Sternberg, 1998, p. 436). According to Sternberg, these puzzles were mainly about determining the architecture of the cognitive mechanisms underlying the task. But the puzzles are truly about how plausible are each model. Now that we know more regarding the source of the fast-same effect, the cognitive architecture and the processing capacity, we can be more critique about the models presented in Chapter 1.

Dual-Process Model

Despite being the predominant view in the 1970s, the dual-process model was abandoned in the 1980s for single-process alternatives. Although the results reported in this thesis do not support the dual-process model, there are a few ideas that it seemingly had right. First, it highlighted that participants likely ponder response alternative serially. Although they do not have a priority for processing similarity, like it was suggested by the model, they do process information one after the other. This behaviour is predicted by the dual-process model, but only for *Different* trials. For *Same* trials, it proposes that participants process information in a more parallel fashion. This view is not supported by the data reported in this thesis.

Bamber himself did not put much emphasis on processing capacity, but some researchers did. For instance, Link and Tindall (1971) were right about the limited capacity nature of the task. Especially in *Different* trials, increasing the workload limited the processing efficiency of participants. In *Same* trials, the processing capacity slows down at a shallower rate than in *Different* trials. This suggests that the processing is dynamic, that it changes over time. Jeff Miller (1978) said that participants are benefiting from redundant information, but attributed this effect to all the trials (not only *Same* trials).

The dual-process model is not a good model of the *Same-Different* task. *Same* and *Different* judgments share similar architectural characteristics and are likely the result of a single processing unit. That said, it is possible to affect the processing efficiency of this processor, and this is what happens in *Same* trials: the context (i.e., a mental representation of the stimulus and the workload) allows the processor to be more efficient and therefore be faster to respond.

Criterion Shift Model

The main contribution of the criterion shift model is that it proposed a single-process alternative to the dual-process model. This model states that participants tend to recheck the elements composing the stimulus to improve the accuracy of the perception. Consequently, participants use a higher decision threshold in the *Different* condition compared to the *Same* condition. Although none of the studies reported in this thesis directly address this question, Study 1 provided some form of evidence against this model. Indeed, by fitting the LBA to the data, I was able to estimate the decision threshold of participants when they responded *All-Same* and *All-Different*. The decision thresholds in these conditions were comparable, despite a large difference in mean response times. This suggests that the decision threshold does not affect the presence of the fast-same effect.

Sensitivity Shift Model

The sensitivity shift model is one of the most influential models of the *Same-Different* task. The model is quite simple: there is only one processor and its processing rate can either be facilitated by the repetition of the same stimuli in a short period of time or inhibited by an incongruence between the stimuli. The Study 1 of this thesis directly tested one critical assumption of this model: the idea that the fast-same effect originates from the stimulus and not from the participant. Indeed, if the fast-same effect is the facilitation of the accumulation rate, we should observe it even if the context of the response is changed; in this case, if the participants press the same answer key to respond *All-Same* and *All-Different*. The results of Study 1 highly support this hypothesis. The

LBA fit also shows support for this model, as it suggests that when all the elements of the stimuli are matching, the accumulation increases compared to when all the elements are mismatching. Further, this is not observed for other parameters of interest, such as response bias, decision threshold and motor response. Therefore, the assumption that the fast-same effect originates from the stimulus and not from the participant is supported by the data.

The extent of the repetition effect remains unclear. Recently, in his doctoral thesis, Harding (2018) showed that facilitation can be modulated by the degree of physical similarity between the stimuli (i.e., manipulating the colour, the font, the case, the phonology). However, in Study 2, the fast-same effect disappeared when the correct response was *Same*, but the letters were physically similar (not identical). The shared physical similarities between two letters should have, according to the sensitivity shift model, facilitated the accumulation rate in favour of the *Same* response, but the results of the experiment refute this hypothesis. Undeniably, there is a repetition effect involved in the fast-same effect, but the cognitive mechanisms explaining this effect are unclear.

Another assumption made by the sensitivity shift model, one that is not explicit, however, is that participants accumulate information in parallel. According to this model, there is either facilitation or inhibition, and it occurs at a constant rate throughout the processing of the stimulus; the model does not have a temporal component. Even though the results of this thesis are not conclusive enough to make any strong claims about the cognitive architecture underlying *Same-Different* judgments, they tend towards seriality. In other words, it is likely that participants ponder response alternatives in succession, process different dimensions of the stimulus in succession and process the elements composing stimuli in succession. If such is the case, facilitation and inhibition might occur at a distinct moment in the process. If anything, this thesis highlights the importance of modelling *Same-Different* judgment using a more dynamic and temporal perspective.

Response Competition

This thesis does not directly address the main aspect of the model, which is the impact of response competition on the fast-same effect. That said, in Study 2, participants were asked to base their judgments partially on the similarity level of the second stimulus compared to the first stimulus. The fast-same effect was observed only when the second stimulus was identical to the first stimulus, not when it was only similar. When the stimulus is similar, but not identical, participants must inhibit the evidence in favour of the *Different* response that occurs because the letters are still mismatching, despite their similarity. In that sense, the response competition model makes sense: accumulating evidence in favour of two response alternatives influence the processing rate of each response.

That said, the model does not explain why participants were still slower to respond *Different* when the letters were dissimilar. Technically, in those trials, there is almost no inhibition for the *Different* accumulator. Likewise, in Study 3, *Different* trials were composed of only mismatching letters, never matching letters. Yet, the fast-same effect was still observed. For this reason, the results reported in this thesis do not support the response competition model.

Response Bias Model

A way to summarize the response bias model is to portray the fast-same effect as the by-product of the human condition. Humans have biases, probably one for overt response such as *Same*, meaning that in the *Same-Different* task, they tend to favour the *Same* response. This has some ecological validity. For example, humans recognize familiar faces more easily (Jung et al., 2013; Schwaninger et al., 2002; Wu et al., 2012) and they better recall information in familiar environments (Godden & Baddeley, 1975). That said, the results of this thesis do not support that the fast-same effect is caused by response bias. In Study 1, even though response bias was removed from the task, participants still responded faster in the *All-Same* condition compared to the *All-Different* condition. The LBA fit also showed comparable response bias parameter values between the two conditions. In Study 2, the fast-same effect disappears when the stimuli are similar (and not identical), even though the response remained *Same*. This shows that response time is not dependent on the response, but truly on the processing rate.

Attention-Driven Single Processor

In the last review article on the *Same-Different* task, Sternberg (1998) barely mentions the attention-driven single processor model. In fact, he simply recommended Farell's (1985) article as supplementary reading. Contrary to the dual-process, the sensitivity shift and the response bias models, Farell’s conceptualization of the *Same-Different* task did not have much of an impact on recent *Same-Different* research. This is regrettable, because the model is probably the best to explain *Same-Different* judgments.

Throughout the experiments reported in this thesis, there is one aspect that seems to constantly come back: attention. It seems that participants would avoid using attentional resources as much as possible. First, they can likely answer using only a pre-attentive treatment of the stimulus. This is especially supported in Study 3, in which the processing capacity of the participants is greater in *Same* trials compared to *Different* trials, but only for the first 500 ms of processing. The capacity of a cognitive system reflects how it manages its processing resources. The system can spend resources on attention to process the stimulus more thoroughly, at the cost of processing speed. When fewer attentional resources are necessary (as it is the case in *Same* trials), the stimulus is processed more rapidly and can even benefit from redundancy in the signal.

The attention-driven single processor also assumes that the processor can use a pre-attentive scan of the stimulus to guide attention. This assumption is supported in Study 2, as participants used a self-terminating stopping rule despite the task requiring an exhaustive treatment of the two diagnostic dimensions (similarity and clarity). Using a pre-attentive scan of the stimulus, participants can gather some information about the global aspect of the stimulus, which then guides how attentional resources are used. This also supposes that the processing of the stimulus is made serially, first using a more global and pre-attentive stage, and second using attention for an analytical processing.

Conclusion and Future Directions

The big answer to the question "*why are humans faster to respond Same than they are to respond Different*" is that they are more efficient to process matching information than mismatching information. The results of the studies presented in this thesis suggest that the cause of the fast-same effect lies in the stimulus, that participants ponder response alternatives sequentially and that participants use pre-attentive scan of stimuli to either respond or guide their attention. Knowing this, the fast-same effect is a consequence of a good management of resources.

Using the capacity coefficient, I was able to compare resource management in *Same* and *Different* trials. The capacity advantage in *Same* trials is not big and rapidly decreases before reaching a limited capacity asymptote at around 500 ms. This time-limited capacity enhancement likely occurs because participants have a mental representation of a *Same* stimulus in working memory. In that case, they can perform one-to-one mapping between the mental representation and the displayed stimulus, a comparison that does not require attention, or at least, not many attentional resources. For the *Different* response, the comparison is many-to-one, that is many potential alternatives to one displayed stimulus. To make such comparison, participants must use more attentional resources.

At the light of these results, future studies on *Same-Different* judgments should focus more on attention. As discussed earlier, most recent research on the topic studied the modulation of the fast-same effect across multiple levels of comparison. If the fast-same effect is caused by a larger processing capacity, it predicts that capacity should also be modulated by the level of comparison. Theoretically speaking, it means that more abstract levels of processing require more attentional resources. This can be tested by estimating the capacity coefficient of participants in distinct tasks using different levels of comparisons. Future research could also focus on the dynamic aspect of the capacity enhancement for *Same* trials using physiological data (i.e., electroencephalographic data). If capacity is indeed enhanced in *Same* trials, it should be reflected in brain activity.

Final Remarks

As final remarks to this thesis, I would like to highlight some aspects that might not seem evident at first glance. Although the literature on *Same-Different* judgments is rich and diversified, it mostly occurred between the mid-1960s and the end of the 1980s. Since that time, psychological research, particularly fundamental cognitive research, has vastly changed. More complex analyses are now easily achievable because computational power has immensely improved in the past few decades. New statistical tools were also developed, such as systems factorial technology, which although it is not directly made to analyze *Same-Different* results, was very useful to answer the research questions of this thesis.

This has forced me to develop creative experimental designs and explore *Same-Different* judgments using a unique perspective. That led to some unexpected results and ultimately newer research questions. It also means that, sometimes, the original research question cannot be answered as strongly as anticipated. For instance, in Study 3, the blurring manipulation unexpectedly did not slow down most participants, yielding any conclusion about the cognitive architecture, ironically, blurry. This is something that I have come to learn over the past few years: data are never clear. The uncertainty of the data is, for me, a great source of stress, mostly because it means that I must decide how to explore the data, and I must include my own degree of freedom in this data analysis process. In this thesis, I learned that exploring the data with post hoc information (i.e., knowing what the data look like) is not only inevitable, but is an important part of scientific progress, as it forced me to consider alternative explanations.

I believe that the only way to be completely serene about this aspect of scientific research is to be as transparent as possible. For this reason, I decided to preregister all the research questions and planned analyses of this thesis. The preregistration forms are all available on the Open Science Framework project pages of the studies. For Study 3, the introduction, methods and analysis plan sections were peer reviewed, and approved by *Attention, Perception, & Psychophysics* prior to data collection. This publication method is very recent and ensure that the article is accepted for publication on the basis of its scientific value, regardless of the direction of the reported effects (Nosek & Lakens, 2014).

Did I follow the analysis plans scrupulously? No. As I mentioned, I took some decisions after seeing the results, I conducted some post hoc analyses, I even changed some planned analyses because some variables interacted in a way that I did not anticipate. That said, those decisions are explicitly mentioned in the articles. My intention was that the reader would know which analyses were confirmatory and exploratory; what decisions were taken before data collection and after data collection.

Also, for sake of transparency and to facilitate future reproductions, I included links to the material used in the study, to raw data and to analysis scripts. Therefore, anyone can access these data and replicate the analyses reported in the articles. Data inaccessibility can be a major hurdle for many researchers who wish to plan their studies accordingly, to systematically review or meta-analyze research findings, or even to calculate statistical power (Alsheikh-Ali et al., 2011; Goulet & Cousineau, 2019b; Wicherts et al., 2006).

Finally, I interpreted the results with as much nuance as possible. Although my experiments were a priori well powered (above the recommended 80% power), I did not want to rely solely on statistical thresholds to interpret the data. This is difficult, especially when using techniques such as systems factorial technology, which rely both on qualitative diagnostic of interaction curves and statistical inferences using null hypothesis statistical testing. However, I also wanted to rely on effect size and tendencies to describe the results. For this reason, the only sentence of this entire thesis that contains the word *significant*, or any of its variation, is this one. The dichotomy of scientific results can lead to misinterpretation, especially when the results are negative (Amrhein et al., 2019). Data are messy, scientific research is not a direct path, results are never black or white, and this thesis is not an exception. But at least, it is transparent.

References

Alsheikh-Ali, A. A., Qureshi, W., Al-Mallah, M. H., & Ioannidis, J. P. A. (2011). Public availability of published research data in High-Impact journals. *PLoS ONE, 6*(9), 2009–2012. https://doi.org/10.1371/journal.pone.0024357

Altieri, N., Fifić, M., Little, D. R., & Yang, C.-T. (2017). Historical foundations and a tutorial introduction to systems factorial technology. In Daniel R Little, N. Altieri, M. Fifić, & C.-T. Yang (Eds.), *Systems factorial technology: A theory driven methodology for the identification of perceptual and cognitive mechanisms* (pp. 3–26). Academic Press. https://doi.org/10.1016/B978-0-12-804315-8.00002-1

Amrhein, V., Greenland, S., & Mcshane, B. (2019). Retire statistical significance. *Nature*, *567*, 305–307. https://doi.org/10.1038/d41586-019-00857-9

Angiolillo-Bent, J. S., & Rips, L. J. (1982). Order information in multiple-element comparison. *Journal of Experimental Psychology. Human Perception and Performance*, *8*(3), 392–406. https://doi.org/10.1037/0096-1523.8.3.392

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*(2), 154–179. https://doi.org/10.1037/0033-295X.93.2.154

Atkinson, R. C., Holmgren, J. E., & Juola, J. F. (1969). Processing time as influenced by the number of elements in a visual display. *Attention, Perception, & Psychophysics*, *6*(6), 321–326. https://doi.org/10.3758/BF03212784

Bagnara, S., Simion, F., & Umiltá. (1984). Reference patterns and the process of normalization. *Perception & Psychophysics*, *35*(2), 186–192. https://doi.org/10.3758/BF03203898

Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for

ANOVA. *Behavior Research Methods*, *44*(1), 158–175. https://doi.org/10.3758/s13428-011-0123-7

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

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

Bamber, D., Herder, J., & Tidd, K. (1975). Reaction times in a task analogous to “same”-"different" judgment. *Perception & Psychophysics*, *18*(5), 321–327. https://doi.org/10.3758/BF03211207

Bamber, D., & Paine, S. (1973). Information retrieval processes in “Same”-"Different" judgments of letter strings. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 477–49). Academic Press.

Belke, E., & Meyer, A. S. (2002). Tracking the time course of multidimensional stimulus discrimination: Analyses of viewing patterns and processing times during “same”-"different" decisions. *European Journal of Cognitive Psychology*, *14*(2), 237–266. https://doi.org/10.1080/09541440143000050

Beller, H. K. (1970). Parallel and serial stages in matching. *Journal of Experimental Psychology*, *84*(2), 213–219. https://doi.org/10.1037/h0029065

Ben-David, B. M., & Algom, D. (2009). Species of redundancy in visual target detection. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(4), 958–976. https://doi.org/10.1037/a0014511

Bindra, D., Donderi, D. C., & Nishisato, S. (1968). Decision latencies of same and different judgments. *Perception & Psychophysics*, *3*(2B), 121–130. https://doi.org/10.3758/BF03212780

Bindra, D., Williams, J. A., & Wise, J. S. (1965). Judgments of sameness and difference : Experiments on decision time. *American Association for the Advancement of Science*, *150*(3703), 1625–1627. https://doi.org/10.1126/science.150.3703.1625

Boles, D. B., & Clifford, J. E. (1989). An upper- and lowercase alphabetic similarity matrix, with derived generation similarity values. *Behavior Research Methods, Instruments, & Computers*, *21*(6), 579–586. https://doi.org/10.3758/BF03210580

Boyerinas, B. M. (2016). Determining the statistical power of the Kolmogorov-Smirnov and Anderson-Darling goodness-of-fit tests via Monte Carlo simulation. In *CNA Occasional Paper* (Issue December).

Brodeur, A., Lé, M., Sangnier, M., & Zylberberg, Y. (2012). Star Wars: The empirics strike back. *SSRN Electronic Journal*, *33*(0), 0–37. https://doi.org/10.2139/ssrn.2089580

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

Buffart, H., & Geissler, H.-G. (1984). Task-dependent representation of categories and memory-guided inference during classification. In E. Degreef & J. van Buggenhaut (Eds.), *Trends in Mathematical Psychology* (pp. 33–58). Elsevier Science Publishers B.V. (North Holland). https://doi.org/10.1016/S0166-4115(08)62079-8

Bundesen, A. C., Larsen, A., Farrell, J. E. J., C, B., Larsen, A., & Farrell, J. E. J. (1981).

Mental transformations of size and orientation. In J. Long & A. Baddeley (Eds.), *Attention and Performance IX* (pp. 279–294). Lawrence Erlbaum Associates.

Bushmakin, M. A., Eidels, A., & Heathcote, A. (2017). Breaking the rules in perceptual information integration. *Cognitive Psychology*, *95*, 1–16. https://doi.org/10.1016/j.cogpsych.2017.03.001

Chechile, R. A. (2011). Properties of reverse hazard functions. *Journal of Mathematical Psychology*, *55*(3), 203–222. https://doi.org/10.1016/j.jmp.2011.03.001

Chignell, M. H., & Krueger, L. E. (1984). Further evidence for priming in perceptual matching: Temporal, not spatial, separation enhances the fast-same effect. *Perception & Psychophysics*, *36*(3), 257–265. https://doi.org/10.3758/BF03206367

Cohen, A. L., & Nosofsky, R. M. (2000). An exemplar-retrieval model of speeded same–different judgments. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1549–1569. https://doi.org/10.1037/0096-1523.26.5.1549

Cousineau, D. (2004). Merging race models and adaptive networks: A parallel race network. *Psychonomic Bulletin & Review*, *11*(5), 807–825. https://doi.org/10.3758/BF03196707

Cousineau, D. (2017). Varieties of confidence intervals. *Advances in Cognitive Psychology*, *13*(2), 140–155. https://doi.org/10.5709/acp-0214-z

Cousineau, D., & Shiffrin, R. M. (2004). Termination of a visual search with large display size effects. *Spatial Vision*, *17*(4–5), 327–352. https://doi.org/10.1163/1568568041920104

Cousineau, D., Thivierge, J. P., Harding, B., & Lacouture, Y. (2016). Constructing a group distribution from individual distributions. *Canadian Journal of Experimental*

Psychology, *70*(3), 253–277. https://doi.org/10.1037/cep0000069

Crist, W. B. (1981). Matching performance and the stimilarity structure of the simulus set. *Journal of Experimental Psychology: General*, *110*(3), 269–296.

Davelaar, E. J., Tian, X., Weidemann, C. T., & Huber, D. E. (2011). A habituation account of change detection in same/different judgments. *Cognitive, Affective, & Behavioral Neuroscience*, *11*(4), 608–626. https://doi.org/10.3758/s13415-011-0056-8

DeCarlo, L. T. (2013). Signal detection models for the same-different task. *Journal of Mathematical Psychology*, *57*(1–2), 43–51. https://doi.org/10.1016/j.jmp.2013.02.002

Decker, L. R. (1974). The effect of method of presentation, set, and stimulus dimensions on "same" - "different" reaction times. *Perception & Psychophysics*, *16*(2), 271–275. https://doi.org/10.3758/BF03203941

Derks, P. L. (1972). Visual recognition of similarity and identity. *Journal of Experimental Psychology*, *95*(1), 237–239. https://doi.org/10.1037/h0033282

Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology. General*, *109*(1), 75–97. https://doi.org/10.1037/0096-3445.109.1.75

Donderi, D. C., & Zelnicker, D. (1969). Parallel processing in visual same-different. *Perception & Psychophysics*, *5*(4), 197–200.

Donders, F. C. (1969). On the speed of mental processes. In K. W.G. (Ed.), *Attention and Performance II* (pp. 412–431). North-Holland Publishing Company. https://doi.org/10.1016/0001-6918(69)90065-1

Donkin, C., Brown, S. D., & Heathcote, A. (2009). The overconstraint of response time models: Rethinking the scaling problem. *Psychonomic Bulletin & Review*, *16*(6), 1129–1135. https://doi.org/10.3758/PBR.16.6.1129

Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E. (2011). Diffusion versus linear ballistic accumulation: different models but the same conclusions about psychological processes? *Psychonomic Bulletin & Review*, *18*, 61–69. https://doi.org/10.3758/s13423-010-0022-4

Donkin, C., Little, D. R., & Houpt, J. W. (2014). Assessing the speed - accuracy trade-off effect on the capacity of information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1183–1202. https://doi.org/10.1037/a0035947

Dyer, F. N. (1973). Same and different judgments for word-color pairs with “irrelevant” words or colors: evidence for word-code comparisons. *Journal of Experimental Psychology*, *98*(1), 102–108. https://doi.org/10.1037/h0034278

Egeth, H. E. (1966). Parallel versus serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics*, *1*(4), 245–252. https://doi.org/10.3758/BF03207389

Eidels, A., Donkin, C., Brown, S. D., & Heathcote, A. (2010). Converging measures of workload capacity. *Psychonomic Bulletin & Review*, *17*(6), 763–771. https://doi.org/10.3758/PBR.17.6.763

Eidels, A., Houpt, J. W., Altieri, N., Pei, L., & Townsend, J. T. (2011). Nice guys finish fast and bad guys finish last: Facilitatory vs. inhibitory interaction in parallel systems. *Journal of Mathematical Psychology*, *55*(2), 176–190. https://doi.org/10.1016/j.jmp.2010.11.003

Entus, A., & Bindra, D. (1970). Common features of the “repetition” and “Same-Different” effects in reaction time experiments. *Perception & Psychophysics*, *7*(3), 143–148. https://doi.org/10.3758/BF03208643

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Eriksen, C. W., O’Hara, W. P., & Eriksen, B. (1982). Response competition effects in same-different judgments. *Perception & Psychophysics*, *32*(3), 261–270. https://doi.org/10.3758/BF03206230

Eviatar, Z., Zaidel, E., & Wickens, T. (1994). Nominal and physical decision criteria in same-different judgments. *Perception & Psychophysics*, *56*(1), 62–72. https://doi.org/10.3758/BF03211691

Farell, B. (1977). *Encoding and comparisons in “Same”-"Different" judgments* [Unpublished doctoral thesis]. McGill University.

Farell, B. (1984). Attention in the processing of complex visual displays: detecting features and their combinations. *Journal of Experimental Psychology. Human Perception and Performance*, *10*(1), 40–64. https://doi.org/10.10370096-1523.10.1.40

Farell, B. (1985). “Same” - “Different” judgments: A review of current controversies in perceptual comparisons. *Psychological Bulletin*, *98*(3), 419–456. https://doi.org/10.1037/0033-2909.98.3.419

Farell, B. (1988). Comparison requirements and attention in identical-nonidentical stimulus discriminations. *Journal of Experimental Psychology. Human Perception* *and Performance*, *14*(4), 707–715. https://doi.org/10.1037/0096-1523.14.4.707

Fifić, M., & Little, D. R. (2017). Stretching mental processes: An overview of and guide for SFT applications. In Daniel R Little, N. Altieri, M. Fifić, & C.-T. Yang (Eds.), *Systems factorial technology: A theory driven methodology for the identification of perceptual and cognitive mechanisms* (pp. 27–52). Academic Press.

Fifić, M., Little, D. R., & Nosofsky, R. M. (2010). Logical-rule models of classification response times: a synthesis of mental-architecture, random-walk, and decision-bound approaches. *Psychological Review*, *117*(2), 309–348. https://doi.org/10.1037/a0018526

Garner, W. R., & Clement, D. E. (1963). Goodness of pattern and pattern uncertainty. *Journal of Verbal Learning and Verbal Behavior*, *2*, 446–452. https://doi.org/10.1016/S0022-5371(63)80046-8

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*(3), 325–331. https://doi.org/10.1111/j.2044-8295.1975.tb01468.x

Godwin, H. J., Walenchok, S. C., Houpt, J. W., Hout, M. C., & Goldinger, S. D. (2015). Faster than the speed of rejection: Object identification processes during visual search for multiple targets. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(4), 1007–1020. https://doi.org/10.1037/xhp0000036

Goulet, M. A., & Cousineau, D. (2019a). Sequential sampling models of Same-Different data and how they explain the fast-same effect. *Canadian Journal of Experimental Psychology*, 1–59. https://doi.org/10.1037/cep0000197

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

Goulet, M. A., & Cousineau, D. (2020). Cognitive architecture and capacity of the cognitive system responsible for Same-Different judgments. *Attention, Perception, & Psychophysics*, 1–18. https://doi.org/10.3758/s13414-020-02008-z

Grice, G. R., Canham, L., & Gwynne, J. W. (1984). Absence of a redundant-signals effect in a reaction time task with divided attention. *Perception & Psychophysics*, *36*(6), 565–570. https://doi.org/10.3758/BF03207517

Griffiths, D. W., Blunden, A. G., & Little, D. R. (2017). Logical-rule based models of categorization: Using systems factorial technology to understand feature and dimensional processing. In *Systems Factorial Technology: A Theory Driven Methodology for the Identification of Perceptual and Cognitive Mechanisms* (1st ed., Issue 2000). Elsevier Inc. https://doi.org/10.1016/B978-0-12-804315-8.00015-X

Grill, D. P. (1971). Variables influencing the mode of processing of complex stimuli. *Perception and Psychophysics*, *10*(1), 51–57. https://doi.org/10.3758/BF03205768

Harding, B. (2018). *A single process model of the Same-Different task* [Unpublished doctoral thesis]. University of Ottawa. https://doi.org/10.20381/ruor-22582

Harding, B., Goulet, M. A., Jolin, S., Tremblay, C., Villeneuve, S.-P., & Durand, G. (2016). Systems factorial technology explained to humans. *The Quantitative Methods for Psychology*, *12*(1), 39–56. https://doi.org/10.20982/tqmp.12.1.p039

Harding, B., LeBlanc, V., Goulet, M. A., & Cousineau, D. (2017). Applying systems factorial technology to discrete accumulators with varying thresholds. In D. R. Little, N. Altieri, M. Fific, & C.-T. Yang (Eds.), *Systems factorial technology: A* *theory driven methodology for the identification of perceptual and cognitive mechanisms* (1st Editio). Elsevier. https://doi.org/10.1016/B978-0-12-804315-8.00016-1

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. https://doi.org/10.20982/tqmp.10.2.p107

Hawkins, H. L. (1969). Parallel processing in complex visual discrimination. *Perception & Psychophysics*, *5*(I), 56–64. https://doi.org/10.3758/BF03210482

Hawkins, H. L., & Shigley, R. H. (1972). Irrelevant information and processing mode in speeded discrimination. *Journal of Experimental Psychology*, *96*(2), 389–395. https://doi.org/10.1037/h0033642

Heathcote, A., Brown, S., & Cousineau, D. (2004). QMPE: estimating Lognormal, Wald, and Weibull RT distributions with a parameter-dependent lower bound. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, *36*(2), 277–290. https://doi.org/10.3758/BF03195574

Heathcote, A., & Hayes, B. (2012). Diffusion versus linear ballistic accumulation: Different models for response time with different conclusions about psychological mechanisms? *Canadian Journal of Experimental Psychology*, *66*(2), 125–136. https://doi.org/10.1037/a0028189

Hock, H. S. (1973). The effects of stimulus structure and familiarity on same-different comparison. *Perception & Psychophysics*, *14*(3), 413–420. https://doi.org/10.3758/BF03211176

Holmgren, J. E., Juola, J. F., & Atkinson, R. C. (1974). Response latency in visual search with redundancy in the visual display. *Perception and Psychophysics*, *16*(1), 123–128. https://doi.org/10.3758/BF03203264

Houpt, J. W., Blaha, L. M., McIntire, J. P., Havig, P. R., & Townsend, J. T. (2014). Systems factorial technology with R. *Behavior Research Methods*, *46*(2), 307–330. https://doi.org/10.3758/s13428-013-0377-3

Houpt, J. W., & Burns, D. M. (2017). Statisatical analyses for systems factorial technology. In Daniel R Little, N. Altieri, M. Fifić, & C.-T. Yang (Eds.), *Systems factorial technology: A theory driven methodology for the identification of perceptual and cognitive mechanisms* (pp. 53–68). Academic Press. https://doi.org/10.1016/B978-0-12-804315-8.00005-7

Houpt, J. W., & Fifić, M. (2017). Adaptive experimental design for systems factorial technology. *Annual Meeting of the Psychonomics Society*.

Houpt, J. W., & Townsend, J. T. (2010). The statistical properties of the Survivor Interaction Contrast. *Journal of Mathematical Psychology*, *54*(5), 446–453. https://doi.org/10.1016/j.jmp.2010.06.006

Houpt, J. W., & Townsend, J. T. (2012). Statistical measures for workload capacity analysis. *Journal of Mathematical Psychology*, *56*(5), 341–355. https://doi.org/10.1016/j.jmp.2012.05.004

Howell, P., & Stockdale, J. E. (1975). Memory and display search in binary classification reaction time. *Perception and Psychophysics*, *18*(6), 379–388.

Hübner, R. (1998). Hemispheric differences in global / local processing revealed by Same- Different judgements. *Visual Cognition*, *5*(4), 457–468.

https://doi.org/10.1080/713756793

Hylan, J. P. (1903). The distribution of attention - 1. *Psychological Review*, *10*(4), 373–403. https://doi.org/10.1037/h0070820

Hyun, J., Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The comparison of visual working memory representations with perceptual inputs. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(4), 1140–1160. https://doi.org/10.1037/a0015019

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), 0696–0701. https://doi.org/10.1371/journal.pmed.0020124

Ioannidis, J. P. A. (2012). Why science is not necessarily self-correcting. *Perspectives on Psychological Science*, *7*(6), 645–654. https://doi.org/10.1177/1745691612464056

Irwin, R. J., Hautus, M. J., & Francis, M. A. (2001). Indices of response bias in the same-different experiment. *Perception & Psychophysics*, *63*(6), 1091–1100. https://doi.org/10.3758/BF03194527

Jacob, J., Breitmeyer, B. G., & Treviño, M. (2013). Tracking the first two seconds: three stages of visual information processing? *Psychonomic Bulletin & Review*, *20*(6), 1114–1119. https://doi.org/10.3758/s13423-013-0482-4

Jung, K., Ruthruff, E., & Gaspelin, N. (2013). Automatic identification of familiar faces. *Attention, Perception & Psychophysics*, *75*(7), 1438–1450. https://doi.org/10.3758/s13414-013-0468-3

Kinoshita, S., & Kaplan, L. (2008). Priming of abstract letter identities in the letter match task. *The Quarterly Journal of Experimental Psychology*, *61*(12), 1873–1885. https://doi.org/10.1080/17470210701781114

Koriat, A., & Norman, J. (1988). Frames and images : Sequential effects in mental rotation. *Cognition*, *14*(1), 93–111.

Koriat, A., & Norman, J. (1989a). Establishing global and local correspondence between successive stimuli: the holistic nature of backward alignment. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *15*(3), 480–494. https://doi.org/10.1037/0278-7393.15.3.480

Koriat, A., & Norman, J. (1989b). Why is word recognition impaired by disorientation while the identification of single letters is not? *Journal of Experimental Psychology. Human Perception and Performance*, *15*(1), 153–163. https://doi.org/10.1037/0096-1523.15.1.153

Koriat, A., Norman, J., & Kimchi, R. (1991). Recognition of rotated letters: extracting invariance across successive and simultaneous stimuli. *Journal of Experimental Psychology. Human Perception and Performance*, *17*(2), 444–457. https://doi.org/10.1037/0096-1523.17.2.444

Krueger, L. E. (1973). Effect of stimulus frequency on speed of “Same”-"Different" judgments. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 497–506). Academic Press.

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

Krueger, L. E. (1983). Probing Proctor’s priming principle : The effect of simultaneous and sequential presentation on Same-Different judgments. *Journal of Experimental Psychology. Learning, Memory and Cogntiontion*, *9*(3), 511–523. https://doi.org/10.1037/0278-7393.9.3.511

Krueger, L. E. (1985). Effect of intermixed foveal and parafoveal presentation on same-different judgements: Evidence for a criterion-inertia model. *Perception & Psychophysics*, *37*(3), 266–271. https://doi.org/10.3758/BF03207574

Krueger, L. E. (1987). Effect of backward masking on same-different judgments. *Perception & Psychophysics*, *41*(4), 375–381. http://www.ncbi.nlm.nih.gov/pubmed/3588235

Krueger, L. E., & Shapiro, R. G. (1979). Letter detection with rapid serial visual presentation: Evidence against word superiority at feature extraction. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(4), 657–673. https://doi.org/10.1037/0096-1523.5.4.657

Krueger, L. E., & Shapiro, R. G. (1981). A reformulation of Proctor's unified theory for matching-task phenomena. *Psychological Review*, *88*(6), 573–581. https://doi.org/10.1037/0033-295X.88.6.573

Kuhn, T. S. (1961). The function of measurement in modern physical science. *Isis*, *52*(2), 161–193. https://doi.org/10.1086/349468

Lachmann, T. (2001). Strategies of coding and processing in a physical same-different task. *Proceedings of the International Society for Psychophysics*, 308–313.

Lachmann, T., & Geissler, H.-G. (2002). Memory search instead of template matching? Representation-guided inference in same-different performance. *Acta Psychologica*, *111*(3), 283–307. https://doi.org/10.1016/S0001-6918(02)00055-0

Link, S. W., & Tindall, A. D. (1971). Speed and accuracy in comparative judgments of line length. *Perception and Psychophysics*, *9*(3), 284–288. https://doi.org/10.3758/BF03212649

Little, Daniel R., Altieri, N., Fifić, M., & Yang, C.-T. (2017). *Systems factorial technology: A theory driven methodology for the identification of perceptual and cognitive mechanisms* (Daniel R Little, N. Altieri, M. Fifić, & C.-T. Yang (eds.)). Academic Press. https://doi.org/10.1016/B978-0-12-804315-8.00024-0

Little, Daniel R., Eidels, A., Houpt, J. W., & Yang, C.-T. (2017). Set size slope still does not distinguish parallel from serial search. *Behavioral and Brain Sciences*, *40*(e145), 32–33. https://doi.org/10.1017/S0140525X16000157

Little, Daniel R., Nosofsky, R. M., & Denton, S. E. (2011). Response-time tests of logical-rule models of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 1–27. https://doi.org/10.1037/a0021330

Luce, D. R. (1986). Memory scanning, visual search, and Same-Different designs. In *Response Times: Their role in inferring elementary mental organization* (pp. 425–455). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195070019.001.0001

Lupker, S. J., Nakayama, M., & Perea, M. (2015). Is there phonologically based priming in the same−different task? Evidence from Japanese−English bilinguals. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(5), 1281–1299. https://doi.org/10.1037/xhp0000087

Miller, J. (1978). Multidimensional same-different judgments: Evidence against independent comparisons of dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(3), 411–422. https://doi.org/10.1037/0096-1523.4.3.411

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals.

Cognitive Psychology, *14*(2), 247–279. https://doi.org/10.1016/0010-0285(82)90010-X

Miller, J. O., & Pachella, R. G. (1973). Locus of the stimulus probability effect. *Journal of Experimental Psychology*, *101*(2), 227–231. https://doi.org/10.1037/h0035214

Millspaugh, J. R. (1978). Effects of array organization on same-different judgments. *Perception & Psychophysics*, *23*(1), 27–35. https://doi.org/10.3758/BF03214291

Møller, P., Köster, E. P., Dijkman, N., De Wijk, R., & Mojet, J. (2012). Same-different reaction times to odors: Some unexpected findings. *Chemosensory Perception*, *5*(2), 158–171. https://doi.org/10.1007/s12078-012-9124-x

Moneer, S., Wang, T., & Little, D. R. (2016). The processing architectures of whole-object features: A logical-rules approach. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(9), 1443–1465. https://doi.org/10.1037/xhp0000227

Morais, J., & Darwin, C. J. (1974). Ear differences for Same-Different reaction times to monaurally presented speech. *Brain and Language*, *1*, 383–390.

Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 520–538. https://doi.org/10.1037/0096-1523.17.2.520

Naikar, N. (1999). Same/different judgements about the direction and colour of apparent-motion stimuli after commissurotomy. *Neuropsychologia*, *37*, 485–493.

Navarro, D. J. (2019). Between the devil and the deep blue sea : Tensions between scientific judgement and statistical model selection. *Computational Brain & Behavior*, *2*, 28–34. https://doi.org/10.1007/s42113-018-0019-z

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–383. https://doi.org/10.1016/0010-0285(77)90012-3

Neisser, U. (1967). *Cognitive Psychology*. Appleton-Century-Crofts.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, *7*, 308–313. https://doi.org/10.1093/comjnl/7.4.308

Nickerson, R. S. (1965). Response times for “Same” “Different” judgements. *Perceptual & Motor Skills*, *20*, 15–18. https://doi.org/10.2466/pms.1965.20.1.15

Nickerson, R. S. (1967a). Categorization time with categories defined by disjunctions and conjunctions of stimulus attributes. *Journal of Experimental Psychology*, *73*(2), 211–219. https://doi.org/10.1037/h0021270

Nickerson, R. S. (1967b). “Same-”Different" response times with multi-attribute stimulus differences. *Perceptual and Motor Skills*, *24*, 543–554. https://doi.org/10.2466/pms.1967.24.2.543

Nickerson, R. S. (1968). Note on “Same”-"Different" response times. *Perceptual & Motor Skills*, *27*, 565–566.

Nickerson, R. S. (1969). “Same’-’different” response times: A model and a preliminary test. In W. G. Koster (Ed.), *Attention and Performance II* (pp. 257–275). North-Holland Publishing Company.

Nickerson, R. S. (1972). Auditory codability and the short-term retention of visual information. *Journal of Experimental Psychology*, *95*(2), 429–436. https://doi.org/10.1037/h0033660

Nickerson, R. S. (1973a). Frequency, recency, and repetition effects on same and different response times. *Journal of Experimental Psychology*, *101*(2), 330–336. https://doi.org/10.1037/h0035232

Nickerson, R. S. (1973b). The use of binary-classification tasks in the study of human information processing: A tutorial survey. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 449–475). Academic Press.

Nickerson, R. S. (1978). On the time it takes to tell things apart. In J. Requin (Ed.), *Attention and Performance VII* (pp. 77–88). Lawrence Erlbaum Associates.

Nickerson, R. S. (1981). Context is important but it does not explain everything: a comment on “Matching performance and the similarity structure of the stimulus set” by Crist. *Journal of Experimental Psychology: General*, *110*(3), 297–302. https://doi.org/10.1037//0096-3445.110.3.297

Nickerson, R. S., & Pew, R. W. (1973). Visual pattern matching: An investigation of some effects of decision task, auditory codability, and spatial correspondence. *Journal of Experimental Psychology*, *98*(1), 36–43. https://doi.org/10.1037/h0034300

Nosek, B. A., & Lakens, D. (2014). Registered Reports: A method to increase the credibility of published results. *Social Psychology*, *45*(3), 137–141. https://doi.org/10.1027/1864-9335/a000192

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*(2), 266–300. https://doi.org/10.1037/0033-295X.104.2.266

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716–aac4716.

https://doi.org/10.1126/science.aac4716

Pachella, R. G., & Miller, J. O. (1976). Stimulus probability and same-different classification. *Perception & Psychophysics*, *19*(1), 29–34. https://doi.org/10.3758/BF03199382

Pan, K., & Eriksen, C. W. (1993). Attentional distribution in the visual field during same-different judgments as assessed by response competition. *Perception & Psychophysics*, *53*(2), 134–144. https://doi.org/10.3758/BF03211723

Petrov, A. A. (2009). Symmetry-based methodology for decision-rule identification in same--different experiments. *Psychonomic Bulletin & Review*, *16*(6), 1011–1025. https://doi.org/10.3758/PBR.16.6.1011

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, *9*(4), 456–479. https://doi.org/10.1097/00004691-199210000-00002

Popper, K. (1959). *The Logic of Scientific Discovery*. Routledge. https://doi.org/https://doi.org/10.1177/000271626032800174

Posner, M. I., & Mitchell, R. F. (1967). Chronometric analysis of classification. *Psychological Review*, *74*(5), 392–409. https://doi.org/10.1037/h0024913

Proctor, R. W. (1981). A unified theory for matching-task phenomena. *Psychological Review*, *88*(4), 291–326. https://doi.org/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. https://doi.org/10.1037/0033-295X.93.4.473

Proctor, R. W., & Healy, A. F. (1985). Order-relevant and order-irrelevant decision rules in multiletter matching. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *11*(3), 519–537. https://doi.org/10.1037/0278-7393.11.3.519

Proctor, R. W., & Healy, A. F. (1987). Task-specific serial position effects in comparisons of multiletter strings. *Perception and Psychophysics*, *42*(2), 180–194. https://doi.org/10.3758/BF03210507

Proctor, R. W., Healy, A. F., & Van Zandt, T. (1991). Same-different judgments of multiletter strings: insensitivity to positional bias and spacing. *Perception & Psychophysics*, *49*(1), 62–72. https://doi.org/10.3758/BF03211617

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), 1981–1982.

Proctor, R. W., & Rao, K. V. (1983a). Reinstating the original principles of Proctor’s unified theory for matching-task phenomena: An evaluation of Krueger and Shapiro’s reformulation. *Psychological Review*, *90*(1), 21–37. https://doi.org/10.1037/0033-295X.90.1.21

Proctor, R. W., & Rao, K. V. (1983b). Evidence that the same-different disparity in letter matching is not attributable to response bias. *Perception & Psychophysics*, *34*(1), 72–76. https://doi.org/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. https://doi.org/10.3758/BF03203923

Psychology Software Tools. (2019). *INFO: E-Prime 2.0 Timing Data [19579]*. https://support.pstnet.com/hc/en-us/articles/229355247-INFO-E-Prime-2-0-Timing-

Data-19579-

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, *88*(6), 552–572. https://doi.org/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. https://doi.org/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. https://doi.org/10.3758/BF03214286

Ratcliff, R., & Hacker, M. J. (1982). On the misguided use of reaction-time differences: A reply to Proctor and Rao (1982). *Perception & Psychophysics*, *31*(6), 603–604. https://doi.org/10.3758/BF03204201

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., McKoon, G., & Verwoerd, M. (1989). A bias interpretation of facilitation in perceptual identification. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *15*(3), 378–387. https://doi.org/10.1037/0278-7393.15.3.378

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*(2), 333–367. https://doi.org/10.1037/0033-295X.111.2.333

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, *107*(2), 358–367. https://doi.org/10.1037//0033-295X.107.2.358

Sagi, E., Gentner, D., & Lovett, A. (2012). What difference reveals about similarity. *Cognitive Science*, *36*(6), 1019–1050. https://doi.org/10.1111/j.1551-6709.2012.01250.x

Savalei, V., & Dunn, E. (2015). Is the call to abandon p-values the red herring of the replicability crisis? *Frontiers in Psychology*, *6*(MAR), 10–13. https://doi.org/10.3389/fpsyg.2015.00245

Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2017). Comparison of the working memory load in N-back and working memory span tasks by means of EEG frequency band power and P300 amplitude. *Frontiers in Human Neuroscience*, *11*(January), 1–19. https://doi.org/10.3389/fnhum.2017.00006

Schwaninger, A., Lobmaier, J. S., & Collishaw, S. M. (2002). Role of featural and configural information in familiar and unfamiliar face recognition. *Biologically Motivated Computer Vision*, 643–650. https://doi.org/10.1007/3-540-36181-2_64

Sekuler, R. W., & Abrams, M. (1968). Visual sameness : A choice time analysis of pattern recognition processes. *Journal of Experimental Psychology*, *77*(2), 232–238. https://doi.org/10.1037/h0025741

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701–703. https://doi.org/10.1126/science.171.3972.701

Silverman, W. P., & Goldberg, S. L. (1975). Further confirmation of same vs. different processing differences. *Perception & Psychophysics*, *17*(2), 189–193.

https://doi.org/10.3758/BF03203884

Simion, F., Bagnara, S., Roncato, S., & Umilta, C. (1982). Transformation processes upon the visual code. *Perception & Psychophysics*, *31*(1), 13–25. https://doi.org/10.3758/BF03206197

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Sinha, N., & Glass, A. L. (2017). Dissociating medial temporal and striatal memory systems with a Same/Different matching task: Evidence for two neural systems in human recognition. *The Journal of General Psychology*, *144*(2), 110–129. https://doi.org/10.1080/00221309.2016.1276044

Smith, G. D., & Ebrahim, S. (2002). Data dredging, bias, or confounding. *British Medical Journal*, *325*, 1437–1438. https://doi.org/10.1136/bmj.325.7378.1437

Smith, P. L., & Ratcliff, R. (2015). Diffusion and random walk processes. *International Encyclopedia of the Social and Behavioral Sciences*, *6*, 395–401. http://www.sciencedirect.com/science/article/pii/B0080430767006203

Snodgrass, J. G. (1972). Matching patterns vs matching digits: The effect of memory dependence and complexity on "same'-'different" reaction times. *Perception & Psychophysics*, *Vol. 11*(5), 341–349. https://doi.org/10.3758/BF03206264

Spieser, L., Servant, M., Hasbroucq, T., & Burle, B. (2017). Beyond decision! Motor contribution to speed – accuracy trade-off in decision-making. *Psychonomic Bulletin and Review*, *24*, 950–956. https://doi.org/10.3758/s13423-016-1172-9

St. James, J. D., & Eriksen, C. W. (1992). Response competition produces a “fast same effect” in same-different judgments. In J. R. Pomerantz & G. R. Lockhead (Eds.), *Uncertainty and structure as psychological concerpts* (pp. 157–168). American Psychological Association. https://doi.org/10.1037/10101-009

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*(3736), 652–654.

Sternberg, S. (1969). The discovery of processing stages: extensions of Donder’s method. In W. G. Koster (Ed.), *Attention and Performance II* (pp. 276–315). North-Holland Publishing Company. https://doi.org/10.1016/0001-6918(69)90055-9

Sternberg, S. (1998). Inferring mental operations from reaction time data : How we compare objects. In D. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science* (2nd Editio, Issue 4, pp. 365–454). MIT Press. https://doi.org/10.1017/CBO9781107415324.004

Storn, R., & Price, K. (1997). Differential evolution – A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, *11*, 341–359. https://doi.org/10.1023/A:1008202821328

Swan, G., & Wyble, B. (2014). The binding pool: A model of shared neural resources for distinct items in visual working memory. *Attention, Perception, & Psychophysics*, *76*(7), 2136–2157. https://doi.org/10.3758/s13414-014-0633-3

T.-Groulx, J., Harding, B., & Cousineau, D. (2019). The EZ diffusion model: An overview with derivation, software and an application to the Same-Different task. *The Quantitative Methods for Psychology*, 1–51.

Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection.

Psychological Review, *61*(6), 401–409. https://doi.org/10.1037/h0058700

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

Taylor, D. A. (1976b). Holistic and analytic processes in the comparison of letters. *Perception & Psychophysics*, *20*(3), 187–190. https://doi.org/10.3758/BF03198599

Taylor, D. A. (1977). Time course of context effects. *Journal of Experimental Psychology: General*, *106*(4), 404–426. https://doi.org/10.1037/0096-3445.106.4.404

Taylor, R. L. (1969). Comparison of short-term memory and visual sensory analysis as sources of information. *Journal of Experimental Psychology*, *81*(3), 515–522.

Thomas, E. A. C., & Ross, B. H. (1980). On appropriate procedures for combining probability distributions within the same family. *Journal of Mathematical Psychology*, *21*(2), 136–152. https://doi.org/10.1016/0022-2496(80)90003-6

Townsend, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, *25*(2), 168–199. https://doi.org/10.1111/j.2044-8317.1972.tb00490.x

Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge University Press.

Townsend, J. T., & Eidels, A. (2011). Workload capacity spaces: A unified methodology for response time measures of efficiency as workload is varied. *Psychonomic Bulletin & Review*, *18*(4), 659–681. https://doi.org/10.3758/s13423-011-0106-9

Townsend, J. T., Houpt, J. W., & Silbert, N. H. (2012). General recognition theory extended to include response times: Predictions for a class of parallel systems. *Journal of Mathematical Psychology*, *56*(6), 476–494. https://doi.org/10.1016/j.jmp.2012.09.001

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, *39*(4), 321–359. https://doi.org/10.1006/jmps.1995.1033

Townsend, J. T., & Wenger, M. J. (2004). The serial-parallel dilemma: a case study in a linkage of theory and method. *Psychonomic Bulletin & Review*, *11*(3), 391–418. https://doi.org/10.3758/BF03196588

Tversky, B. (1969). Pictorial and verbal encoding in a short-term memory task. *Perception & Psychophysics*, *6*(4), 225–233. https://doi.org/10.3758/BF03207022

van der Heijden, A. H. C., Schreuder, R., Maris, L., & Neerincx, M. (1984). Some evidence for correlated separate activation in a simple letter-detection task. *Perception and Psychophysics*, *36*(6), 577–585. https://doi.org/10.3758/BF03207519

Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review*, *7*(2), 208–256. papers2://publication/uuid/C7B49B0E-BB75-49CA-B00F-B4AB36FAE669

Voss, A., & Voss, J. (2007). Fast-dm : A free program for efficient diffusion model analysis. *Behavior Research Methods*, *39*(4), 767–775. https://doi.org/10.3758/BF03192967

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Smira, M., Epskamp, S., Etz, A., Matzke, D., … Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin and Review*, *25*, 58–76. https://doi.org/10.3758/s13423-017-1323-7

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 and Review*, *14*(1), 3–22. https://doi.org/10.3758/BF03194023

Walker, J. A., & Cousineau, D. (2019). Into the Mind's Eye: Exploring the Fast-Same Effect in the Same-Different Task. *American Journal of Psychology*, *4*, 421–437. https://doi.org/10.5406/amerjpsyc.132.4.0421

Well, A. D., & Green, J. (1972). Effects of color differences in a letter matching task. *Psychonmic Science*, *29*(2), 109–110. https://doi.org/10.3758/BF03336585

Well, A. D., Pollatsek, A., & Schindler, R. M. (1975). Facilitation of Both Same and Different Judgments of Letter Strings by Familiarity of Letter Sequence. *Perception & Psychophysics*, *17*(5), 511–520. https://doi.org/10.3758/BF03203303

Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *American Psychologist*, *61*, 726–728. https://doi.org/10.1037/0003-066X.61.7.726

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, *9*(1), 33–39. https://doi.org/10.1111/1467-9280.00006

Wu, E. X. W., Laeng, B., & Magnussen, S. (2012). Through the eyes of the own-race bias: Eye-tracking and pupillometry during face recognition. *Social Neuroscience*, *7*(2), 202–216. https://doi.org/10.1080/17470919.2011.596946

Wyble, B., Potter, M. C., Bowman, H., & Nieuwenstein, M. (2011). Attentional episodes in visual perception. *Journal of Experimental Psychology: General*, *140*(3), 488–505. https://doi.org/10.1037/a0023612

Yang, C.-T., Altieri, N., & Little, D. R. (2018). An examination of parallel versus coactive processing accounts of redundant-target audiovisual signal processing. *Journal of Mathematical Psychology*, *82*, 138–158. https://doi.org/10.1016/j.jmp.2017.09.003