ISPA

INSTITUTO UNIVERSITÁRIO
CIÊNCIAS PSICOLÓGICAS, SOCIAIS E DA VIDA

STEREOTYPES-AS-CONTEXT IN VISUAL CHANGE DETECTION TASKS

JOÃO DAVID MARTINS

Dissertation Advisor:
Professor Teresa Garcia-Marques, Ph.D.

Dissertation Seminar Coordinator:
Professor Teresa Garcia-Marques, Ph.D.

Dissertation Submitted as Partial Requirement for the Degree of:
MASTER IN PSYCHOLOGY
Specialization in Social and Organizational Psychology

2018

Master's dissertation conducted under the orientation of Professor Teresa Garcia-Marques, Ph.D., presented at ISPA - University Institute for the degree of Master in the specialty of Social and Organizational Psychology.

Acknowledgments

Developing this dissertation was what it ideally is supposed to be: a very stimulating period of learning new things, acquiring valuable skills, and engaging in discussions with others whose reasoning I admire. Contrary to common discourse regarding dissertations, this was truly a fun experience. And while learning is always of itself fun and exciting, such a positive overall experience is only possible in the company of excellent individuals, to whom I owe the utmost gratitude.

To my advisor, Professor Teresa Garcia-Marques, for sharing vital knowledge that significantly enriched this project; for continued support and availability, while, simultaneously, encouraging autonomy and self-reliance; but most of all, for the unfailing example of curiosity, persistence, rigor, and research ethics.

To my internship advisor, Professor Rui Bártolo-Ribeiro, for understanding the personal significance of this project, and providing every condition necessary for me to make it a priority.

Other teachers have also, throughout my years in ISPA, contributed to the success of this experience, to get to where I am now, and to where I want to be in the future. Namely, Professor Maria João Gouveia, whose classes were lessons in critical thinking and enthusiasm for science; and Professor Bruno Soares Rodrigues, who sparked my interest for, and enjoyment of, statistical methods – something that, I am told, is not easy to convince students of.

To my boss, Carlos Sousa, who provided all the flexibility possible in the workplace, allowing me to dedicate my time and attention to academic pursuits, sometimes at the cost of his own time and efforts.

I would like to show my appreciation for the developers of Google's algorithms, namely in the form of Google Scholar, for organizing the vastness of the internet and saving me incalculable hours, while simultaneously giving me a sincere appreciation of all scholars before me who did their amazing work before the advent of the digital age and its conveniences.

Finally, thank you to those who are closest to me, for remaining close to me when all I could think and talk about was this dissertation. I understand now that ANOVAs are not as exciting to you as they are to me. Sorry.

Resumo

Processos de identificação e de detecção visual demonstram diferente sensibilidade às características do contexto. No primeiro caso, estímulos congruentes com o contexto são identificados mais rapidamente do que os incongruentes; no segundo, a incongruência é detectada mais rapidamente. Em qualquer dos casos é clara a influência das expectativas criadas pelo contexto.

Estudos que usam estereótipos enquanto contexto também observam identificação mais rápida para estímulos congruentes com o contexto. Mas pouco se sabe sobre os efeitos deste tipo de contexto em tarefas de detecção. Para testar este efeito utilizou-se uma *flicker task* na qual 72 participantes detectaram a mudança de um objecto em cenas naturais contendo uma mulher ou homem. O objecto, estímulo crítico, foi congruente (e.g., homem-charuto) ou incongruente com a pessoa (e.g., mulher-gravata). Esperava-se diferenças entre estímulos congruentes e incongruentes, com detecções mais rápidas de estímulos incongruentes quando comparados com congruentes. Os resultados mostram que o nível de congruência não tem impacto na velocidade de detecção. Esta foi apenas afectada pelo tempo decorrido entre o início de cada *trial* e a ocorrência de mudança (*threshold*): em *trials* em que esta ocorre aos 1200ms, a detecção é significativamente mais lenta do que quando ocorre aos 400ms ou 800ms.

Discutem-se estes resultados à luz da literatura sobre detecção de mudança e estereótipos, focando problemas metodológicos da tarefa e a possibilidade de que estereótipos possam influenciar tarefas de identificação e detecção diferentemente. O efeito de *threshold* é discutido à luz de uma mudança de estratégia, de procura global e geral, para localizada.

Palavras-chave: estereótipos, percepção, detecção, *flicker task*

Abstract

Visual identification and detection processes show different sensitivity to the same context characteristics. In one case, context-congruent stimuli seem to be identified faster than incongruent stimuli; conversely, incongruency is detected faster than congruency. In both cases, the influence of expectations created from context is clear.

Studies that use stereotypes as context have also shown that stereotype-congruent stimuli are identified faster. But little is known regarding the effect of these contexts in detection tasks. To test this effect, a flicker task was used in which 72 participants detected a changing object in natural scenes containing either a woman or a man. The object, the critical stimulus, was either stereotypically congruent (e.g., man-cigar) or incongruent (e.g., woman-tie). We expected differences between congruent and incongruent stimuli, such that changes in incongruent stimuli would be detected fast than changes in congruent stimuli. Results show that level of congruency did not impact detection speed. Performance was only affected by time between the start of a trial and occurrence of change (threshold): in trials when change happens at 1200ms, detection is significantly slower than in trials when it happens at 400ms or 800ms.

We discuss these results within the body of knowledge concerning change detection and stereotypes, focusing both on methodologic problems with the task and the possibility that stereotypes might influence detection and identification tasks differently. The observed threshold effect is discussed within the framework of a change of strategy, between a gist-based global visual search, and local search.

Keywords: stereotypes, perception, detection, flicker task

Index

List of Tables

List of Figures

List of Appendices

Introduction

Change detection, or its counterpart and more popular term *change blindness*, is a part of our everyday experience. Both detecting and missing a change in a visual scene are part of the regular working of our cognitive processing: some changes are important to detect (e.g., an unusual spot on an x-ray), while others, in the name of cognitive efficiency, are adequately kept under our radar (e.g., ads on a website). This awareness that some changes are important and should be noticed allowed the phenomenon of change blindness to grow out of the field of research into the public's eye (with popular YouTube videos confirming this popularity) and even government policy – namely, the UK Government's Transport For London, who recreated a famous paradigm (Simons & Chabris, 1999) in a video aimed at alerting London drivers that yes, the "I didn't see the cyclist" report is literal, and that attention needs to be directed at that often unnoticed change.

On the road, change blindness is, then, a matter of public safety. It is also common knowledge that, while driving, reaction times are critical, often marking the difference between light, serious, or deadly outcomes of an accident. As such, research that uncovers the conditions in which these reaction times fluctuate is a must-have basis for any informed and effective intervention or policy aimed at reducing the negative outcomes of change blindness. Such interventions are aimed at people, and will likely opt for stimuli containing people – something we are naturally drawn to (Bracco & Chiorri, 2009). Thus, considering the complex ways in which we tend to represent social reality, it is surprising that no previous research has focused on this type of stimuli.

The study presented in this paper aims, then, to explore and clarify how stereotypes influence stimuli detection in a change blindness paradigm. It builds upon previous research, in which two main findings have been reported: a) activating stereotypes facilitates identification of stereotype-congruent stimuli; b) in detection tasks, object-context incongruency facilitates detection. In the sense that the first refers to an identification task with stereotype activation, and the second is a change blindness detection task with no stereotypical context, these findings may be reporting two different psychological mechanisms that, although promoting opposite effects, do not necessarily conflict. Thus, the question here addressed: will stereotypes-as-context benefit a stereotype-congruent or a stereotype-incongruent stimulus in a detection task?

Change Blindness

Change blindness has been repeatedly defined as the difficulty in detecting changes in a visual scene – even those considered large or substantial (Simons, 2000; Simons & Rensink, 2005). Visual changes occur under many forms, some of which have been demonstrated in a laboratory environment – from eye saccades to blinking, jump cuts in video, alternating screens (for a review, see Simons & Levin, 1997) – and some in real-life scenarios, where half of the participants fail to notice when the person they were interacting with is replaced by another actor (Simons & Levin, 1998).

Simons and Levin's review (1997) underlines a crucial aspect of the change blindness phenomenon: the role of expectations. Because our encoding of visual scenes is not verbatim (instead, an abstraction), expectations would lead to better performance in change detection for "schema-inconsistent objects than the schema-consistent ones" (p. 264).

More recently, LaPointe, Lupianez, and Milliken (2013) have emphasized the same effect, as well as its reversal: in *detection* tasks, context-incongruent targets are detected faster; in *identification* tasks, it is the context-congruent targets that are benefited. They proposed a dual process account of these effects: on one hand, a congruency benefit in identification tasks, in which prior knowledge, expertise, and scene schematics play a role; on the other hand, incongruency benefits in detection tasks depend on *attention capture processes* (ACPs), which are sensitive to target-context inconsistencies (see also LaPointe & Milliken, 2016, 2017).

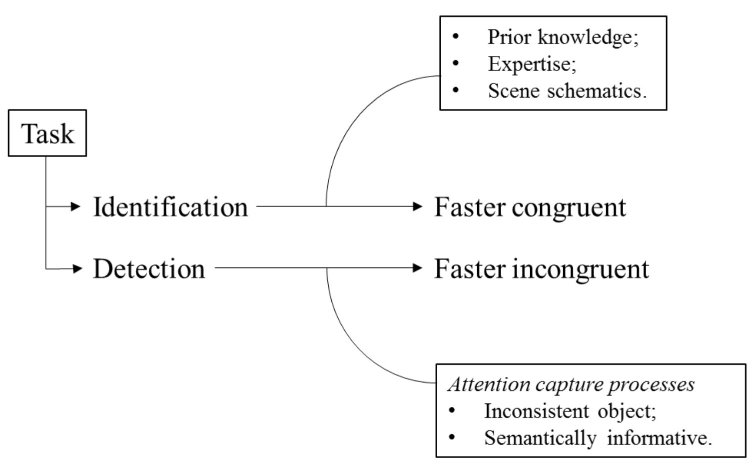


Figure 1. Dual process account of congruency and incongruency benefits, in identification and detection tasks (LaPointe et al., 2013)

This dual process proposal converges with findings that detection and identification tasks rely on different perceptual and neural processes (Busch, Fründ, & Herrmann, 2010; Lyyra, Wikgren, & Astikainen, 2010) and distinct cortical mechanisms (Straube & Fahle, 2011).

Stereotypes and stereotype effects

Picking up on Simons and Levin's (1997) mention of schemas, one can ask: would these congruency- and incongruency-benefits also be observed if the context in which the target stimuli changes contained stereotype-relevant information (i.e., aspects that lead to stereotype activation)? Stereotypes, while defined as structured sets of beliefs and expectations regarding social groups and their members (for a comprehensive definition, see Greenwald & Banaji, 1995), can be paralleled to schemas – both being adaptive and "well-rehearsed, automatically activated cognitive structures" (Cox, Abramson, Devine, & Hollon, 2012, p. 429). As such, stereotypes would act as energy-saving devices, allowing for a more economical and efficient cognition. These were, in fact, the findings of a study by Macrae, Milne, and Bodenhausen (1994): stereotypes allow for less resource-intensive processing, with stereotype-consistent traits being recalled at a higher rate for subjects previously exposed to a matching stereotypic label. In other words, authors observed an expectancy-congruency benefit, both at the conscious and subliminal level.

Recall that stereotypes have a pervasive and diversified influence in our everyday lives, from mediatic race discrimination to academic performance (Cadinu, Maass, Frigerio, Impagliazzo & Latinotti, 2013). With particular relevance to this study, stereotypes have been shown to have subtle and implicit effects on perception, and visual processing (e.g. Eberhardt, Goff, Purdie, & Davies, 2004; Stone, Perry, & Darley, 1997). In short, we know that activating stereotypes influences subsequent processing (e.g., facilitating stimuli identification, differentiated response for congruency and incongruency).

Given this review, and having established that stereotypes can be equated with schemas, we reason that stereotypes can also have a measurable effect on change detection and identification tasks.

While the stereotype-focused studies mentioned above weren't centered on change blindness tasks (or any variation of such), two lines of research regarding *identification* are worth mentioning.

First, research developed within Correll, Park, Judd, and Wittenbrink's (2002) "Police Officer's Dilemma" task: a shoot/don't shoot computerized decision task in which targets (African American vs. White) would appear in a naturalistic background, holding either a weapon or common objects (e.g., cell phone, soda cans, etc.) Participants are instructed to "shoot" when the target is holding a gun, and to press "don't shoot" when they are holding anything else – they have under a second to decide. Notice that race or ethnicity are not relevant to perform this task. The set of studies developed with this task are based on the contents of the African American stereotype, which is associated with violence and danger, and this stereotypic schema is assumed to interfere with the task, biasing identification in favor of stereotype-congruent targets. Results support this interpretation: participants shoot an armed target faster when African American (vs. White) and chose not to shoot an unarmed target faster when White (vs. African American). Error rates show the same tendency.

A second relevant line of research was developed within Payne's (2001) weapon identification task, in which participants, along several trials, have to quickly identify either a gun or a tool, after being primed with black or white faces. Much like in Correll et al. (2002), both the analysis of reaction times and accuracy suggested a stereotype-congruency benefit.

These studies are relevant to our goal because a) they use stereotypes as context; and b) they apply variants of an identification task. In short, they suggest that: much like non-stereotypical schemas benefit the identification of schema-congruent stimuli, stereotypes also benefit the identification of stereotype-congruent stimuli. While these studies present priming, associative networks, and concept activation as mechanisms on which perception may rely in order to benefit context-congruent targets, they follow an *identification* paradigm.

Present study

Our remaining question is if the incongruency-benefit observed in *detection* tasks can also be observed when stereotypes (or stereotype activating stimulus) are the context.

Regarding change detection and identification, the literature reviewed above (and more thoroughly in Appendix A) establishes that:

- In *identification* tasks, performance is benefited for context-congruent targets, by means of prior knowledge, expertise, gist-based scene schematics;

- In *detection* tasks, performance is benefited for context-incongruent targets, by means of attention capture processes, sensitive to context-target inconsistencies.

Regarding stereotypes, however, literature tells us that:

- *Identification* of stereotype-congruent targets is also facilitated, with faster response and lower error rates;

Detection (with stereotypes as context) is hypothesized to, not unlike change blindness, mobilize attention processes that benefit incongruity. This is, in fact, one of the advantages of relying on schemas on our cognition: the quick detection of the unexpected, that which violates assumptions and expectations, and most likely needs our attention. If stereotypes can be equated to other schemas, incongruency should also be benefited in a detection task.

This will be tested in a computerized flicker task (Simons, 2000; Simons & Rensink, 2005, p. 16), in which “an original and modified scene alternate repeatedly, separated by a brief blank display, until observers find the change”. Trials will either be congruent (e.g., a man figures in the context, and the target is an object highly associated with men, such as a cigar) or incongruent (e.g., a man figures in the background, and the target is lipstick).

Every trial was categorized according to three main variables, for which reaction times (henceforth RT) were calculated:

a) Stereotype: whether the person present in the scene was a *woman* or a *man*;
b) Object: determines whether the object that appeared (i.e. the change) was categorized in pre-test as *male* or *female*;
c) Threshold: the time at which the change occurs, varying between 400, 800, or 1200ms.

Congruency is defined as the interaction of stereotype and object gender.

To test our hypothesis, that stereotype-object congruency will impact detection performance as expressed in differing RTs, a flicker task was adapted from Rensink, O’Regan, and Clark (1997). This paradigm, being the most widely used in literature on the topic, has been tested by multiple independent researchers, allowing for replications, comparability of results, and overall convergence of evidence.

Method

Participants

A sample of 74 first and second-year undergraduate students (56 women, 16 men) with ages between 18 and 51 (M = 23, SD = 7.8) was recruited in exchange for credit in two courses at ISPA – University Institute, in Lisbon, Portugal. Sample size was determined in power analysis, using the software G*Power (a priori computation for a repeated measures ANOVA within factors; Faul, Erdfelder, Lang, & Buchner, 2007). As effect sizes in social psychology usually vary between small and medium (Richard, Bond, & Stokes-Zoota, 2003), a conservative effect size was selected in the computation. For $\eta p^2 = 0.1$, a required sample size was 73 participants.

Design

The design is a 2 (Stereotype: Woman vs. Man) x 2 (Object gender: Female vs. Male) x 3 (Threshold: 400ms vs. 800ms vs. 1200ms) design, with all three measures being within-subjects.

Materials

Photographs of objects used in this study were pre-tested (N= 33; 23 female; age = 35,55; SD = 13,39) to ensure their stereotypic fit vs. no fit to Male and Female gender stereotype (see Appendix B). These objects were then superimposed in a photograph of a natural environment. Objects were shown along with a person of the same or different gender from that with which it was matched in the pre-test, hence creating congruent and incongruent trials. To make sure that congruency levels varied only regarding stereotype, the environments were carefully selected so that: a) both a woman and a man could naturally be found in each scene; and b) objects were also naturally found in each scene (e.g., both a woman, a man, a dress, and a tie can be found in a bedroom).

Photos for nine scenes were taken purposely for the study, using a Sony Alpha6000 camera, edited with Adobe Lightroom CC, and composed in Adobe Photoshop CC for Windows. For the remaining scenes, images were obtained online, through a Google Image search, filtering for copyright-free images only. Finally, pictures of objects were either photographed specifically for the study, taken from Javadi and Wee (2012; used with permission), or through a Google Image search, filtering for copyright free images only. All scenes are listed in Appendix C)

From these materials, a total of 72 photographs were used in critical trials – 12 scenes, each with 6 conditions: 2 (stereotype: woman vs. man) x 3 (no object; congruent; incongruent).

Figure 2. Example of a scene; man-congruent condition on the left, and woman-congruent condition on the right.

Procedure

Participants arrived in groups in the laboratory and were welcomed. They were briefed by the experimenter about the general purpose of the experiment and were reminded of the importance of not only their contribution, but also of performing the task with attention. An experimenter was present to answer any question. Afterwards, participants were seated on individual booths and presented with on-screen instructions (see Appendix D).

The task follows the flicker task structure, illustrated in Figure 3 below. For each trial, a fixation crosshair is initially presented for 500ms, followed by an image of a woman or man in a naturalistic scene (200ms). Afterwards, the screen is blank for 200ms, followed by the same image with the critical object (200ms). A blank screen is once more presented for 200ms, thus ending the cycle. This cycle is repeated 19 times or stops when participants detect the change. In both cases, the task advances to the next trial.

To detect random responding, three conditions were created, varying in how many times the first image (A) is shown before the altered image (B). The cycle depicted in Figure 3 shows the 400ms threshold condition, with a single presentation of A before the change occurs. The 800ms condition has two presentations of A, and the 1200ms condition has three presentations of A (always followed by a blank screen). RTs were measured from the moment the change was introduced.

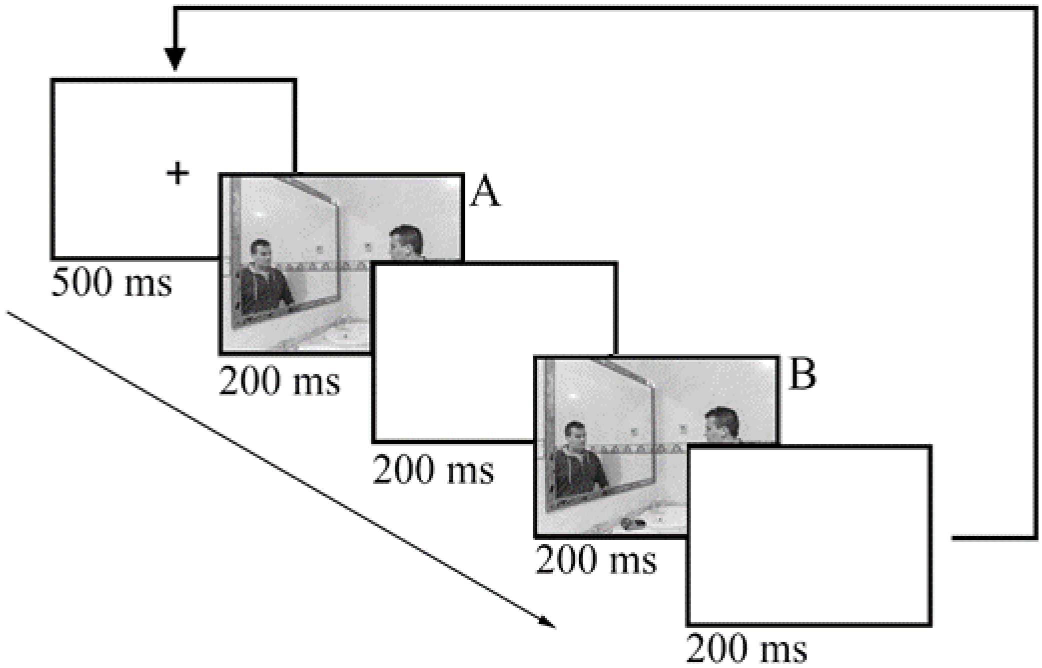


Figure 3. Flicker task, specifically the 400ms condition

Participants were instructed to press SPACEBAR if they detected a change. Instructions clearly indicated *detection*.

In each scene, the changing object was superimposed in the same location. To avoid a learning effect, critical trials were interspersed with lures – trials using the same scene, with changes occurring at different locations, with new non-critical objects or a different category of change (e.g., disappearing light switch; changing the color of an object).

Upon completion of 48 trials, participants were thanked for their contribution.

Results

The main goal of this study was to test if congruency between the gender of a changing object and an exemplar of the gender stereotype in the background would interfere with the speed at which participants detected the change in a visual task. To test this hypothesis, we analyzed participant RTs for each stereotype, object gender, and threshold, as within-participant factors of an ANOVA with the following design: 2 (woman vs. man) x 2 (female vs. male) x 3 (400ms, 800ms, 1200ms).

RT analysis

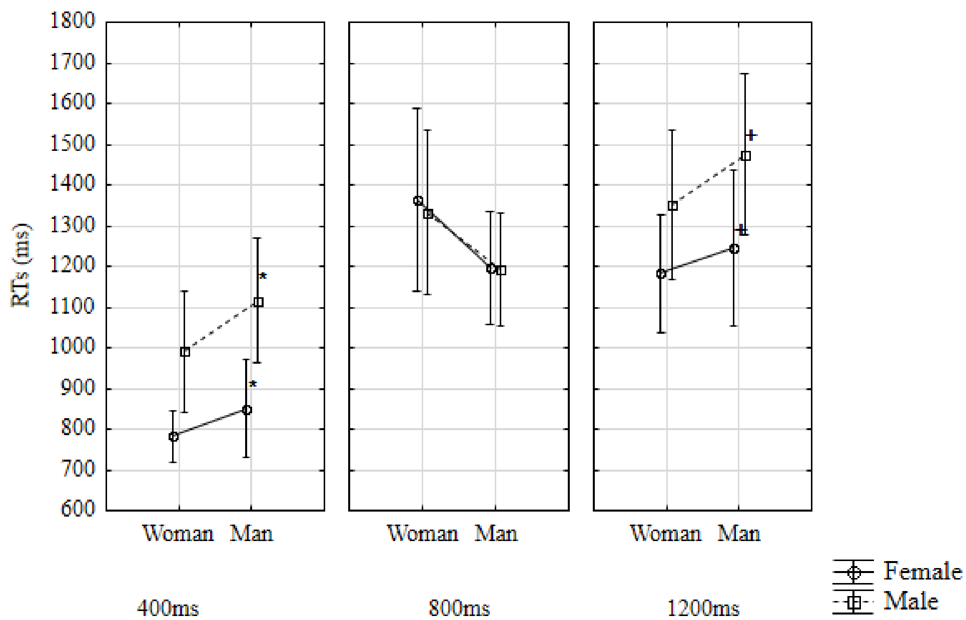


Figure 4. Participant's change detection RTs (ANOVA Stereotype x Object gender x Threshold); * and + indicate significant differences.

Contrary to expectations, no effect for congruency was found, as determined by a non-significant interaction between stereotype and object gender (F (1,71) = .519, p = .474, ηp^2 = .007). Main effects were however detected for: object gender (F (1,71) = 10.669, p = .002, ηp^2 = .131), with objects associated with women being detected faster (M = 1103.04, SD = 39.97) than those associated with men (M = 1243.16, SD = 45.45); and for threshold (F (2,142) = 24.431, p = .000, ηp^2 = .256), which also interacted with other factors, as detailed below.

Figure 5 shows trials with a threshold of 400ms registering significantly lower reaction times when compared to trials with thresholds of 800ms and 1200ms (Fisher's LSD post-hoc test, $p < .05$).

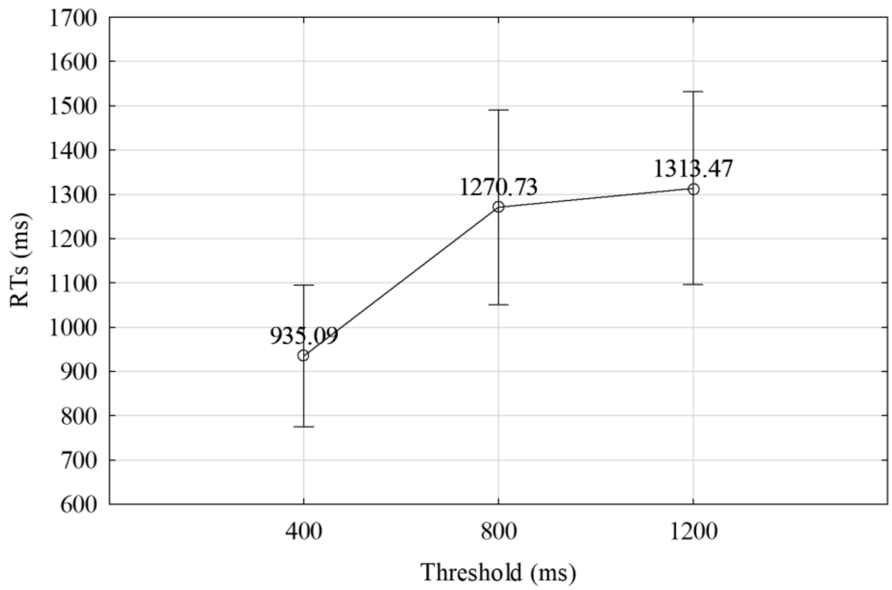


Figure 5. Participant RTs for the three threshold levels

A significant interaction between object gender and threshold was also observed. This effect varied according to threshold level (F (2, 142) = 2.999, p = .053, ηp^2 = .041; Figure 6), with significant differences only at the extremes: 400ms (M_{female} = 816.39, SD_{female} = 37.13 vs. M_{male} = 1053.8, SD_{male} = 60.21) and 1200ms (M_{female} = 1213.57, SD_{female} = 63.46 vs. M_{male} = 1413.38, SD_{male} = 77.51).

A significant interaction between stereotype and threshold was also observed (F (2, 142) = 3.743, p = .026, ηp^2 = .05). At 800ms, trials figuring a man (M = 1193.79, SD = 56.31) were detected significantly faster than those with a woman (M = 1347.67, SD = 81.01; Fisher's LSD, p = .039).

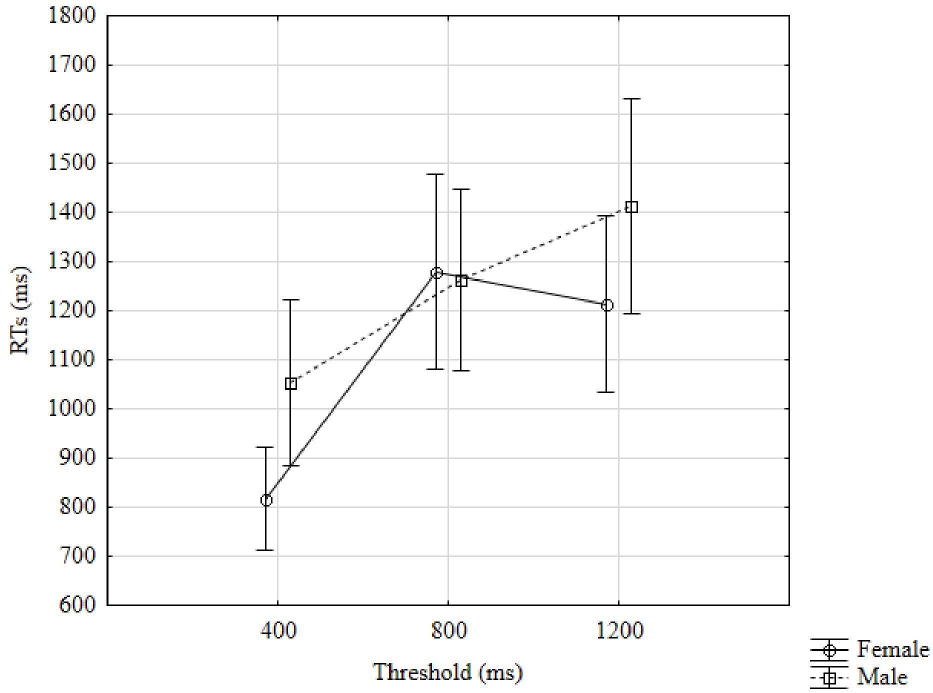


Figure 6. RTs for different object genders by threshold level (ANOVA Object gender vs. Threshold)

In short, we observe that:

a) when change occurs at 400ms and 1200ms, there is a benefit (in the form of lower RTs) for trials containing a female object, regardless of stereotype;
b) when change occurs at 800ms, RTs are lower for trials depicting a man, regardless of object.

These results suggest that differences in detection RTs are not, as expected, due to cognitive processes in which congruency would be considered. Unexpectedly, we found interesting effects of threshold and object gender tied to the materials used to represent the specific object and stereotypes. To obtain a better understanding of the object gender effect, we advance two possibilities that can be addressed with our current data.

The first hypothesis is that the effect is driven by motivational factors. Female and male participants may have been differently motivated to attend to stimuli associated with their own gender. Since most of our participants were women, it is possible that they have higher familiarity with objects typically associated with women than with those associated with men.

Our second hypothesis relates to the perceptive features of the materials used, which could lead to an attention preference benefiting female objects differently from male objects, as well as for the preference for images featuring a man (vs. those featuring a woman) observed in trials with a threshold of 800ms.

Both these post-hoc hypotheses are detailed below, addressing the effects of both object gender and stereotype. Since each level of object gender and stereotype was counterbalanced between each level of threshold, the main effect of threshold is not addressed.

Motivational hypothesis

To test for an effect of a match between participant gender and object gender, the former would have to be introduced as a factor in our previous analysis. Such analysis is not possible due to the uneven distribution of gender in our sample. As such, we performed separate analyses for each gender, keeping in mind that only the female sample analysis has enough power to support a conclusion.

Female sample analysis (n = 55)

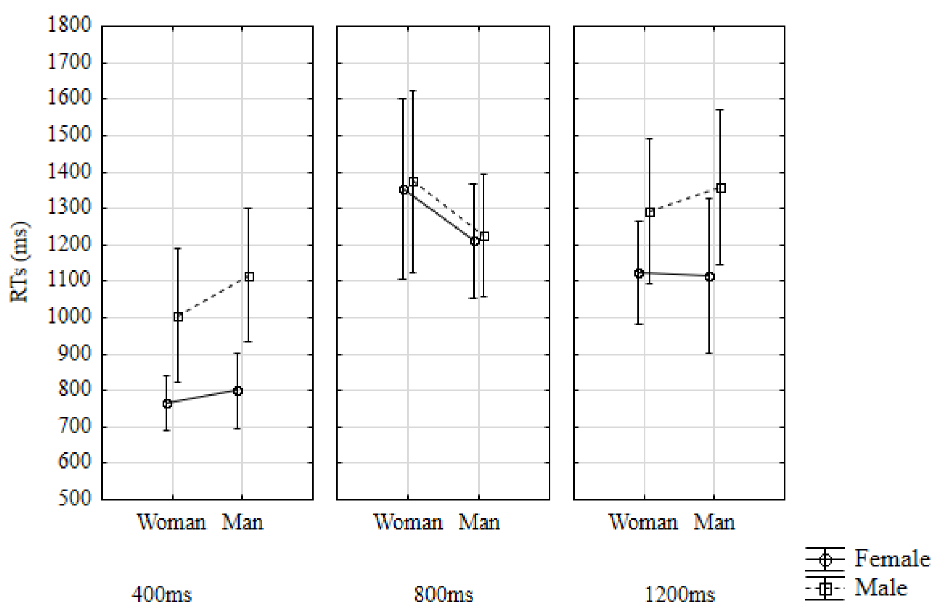


Figure 7. Female participants' RTs (ANOVA Stereotype x Object gender x Threshold)

A 2 (stereotype) x 2 (object gender) x 3 (threshold) ANOVA for female participants' RTs was conducted; replicating previous analyses, no congruency effect, as defined by the interaction between object gender and stereotype, was found (F (1, 54) = .363, p = .549, ηp^2 = .007).

We replicated the effects of object gender (F (1, 54) = 12.195, p = .001, ηp^2 = .184), with female objects being generally detected faster than male objects (M_{female} = 1061.49, SD_{female} = 40.67 vs. M_{male} = 1229.15, SD_{male} = 53.64), and of threshold (F (2, 108) = 19.319, p = .000, ηp^2 = .263), with RTs at 400ms (M = 922.005, SD = 44.9) significantly lower than at 800ms (M = 1291.09, SD = 61.07) and 1200ms (M = 1222.88, SD = 57.58).

Like in the previous analysis, an interaction between object gender and threshold was found, showing that the difference in RTs for object gender is only significant at the outermost levels: 400ms and 1200ms (Fisher's LSD post-hoc test; p = .002).

Interestingly, the interaction between stereotype and threshold is not significant, suggesting that female participants attend to the person in the background equally throughout all levels of threshold.

Male sample analysis (n = 16)

A 2 (stereotype) x 2 (object gender) x 3 (threshold) ANOVA for male participants' RTs was conducted (F (2, 30) = .366, p < .697, ηp^2 = .024). Again, no congruency effect was detected (F (2, 30) = .363, p < .556, ηp^2 = .024).

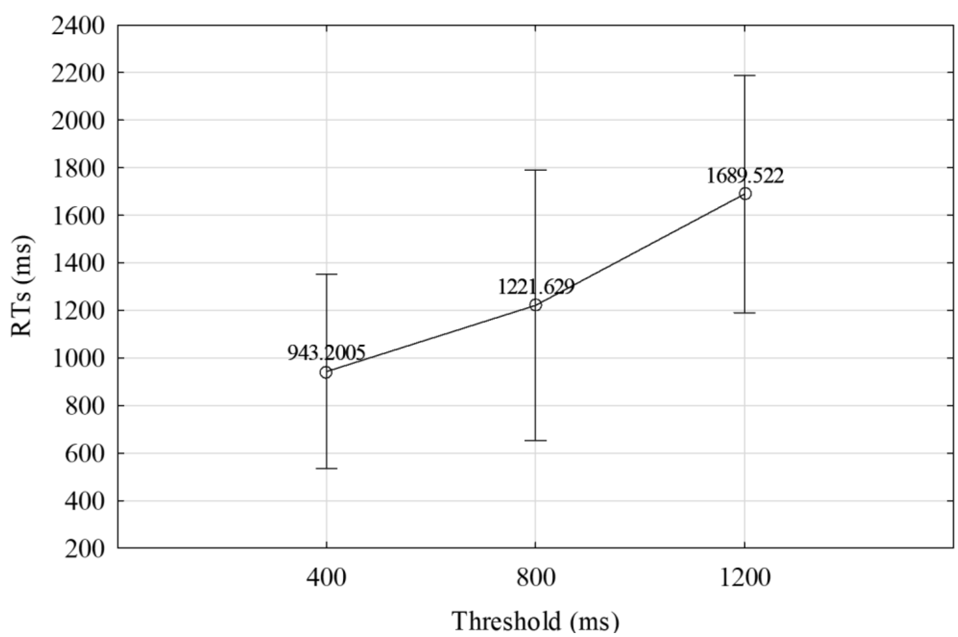


Figure 8. Male participant's RTs for all three levels of threshold.

Not surprisingly, given the reduced sample size, less significant effects were observed with this sample. However, the analysis could be a suggestion that, for this sample, object gender is not relevant, as no effect was detected (F (1, 15) = 0.156, $p < .698$, $\eta p^2 = .01$), with a descriptive power of 0.06.

The only main effect that remained significant was that of threshold (F (2, 30) = 15.39, $p < .001$, $\eta p^2 = .506$; Figure 8); interactions between threshold and stereotype (F (2, 30) = 2.207, $p = .128$, $\eta p^2 = .129$), and threshold and object gender (F (2, 30) = 1.846, $p = .175$, $\eta p^2 = .110$) were not significant.

When analyzing the pattern of means for threshold levels we find that, unlike both the full sample and the female-only sample, each level of threshold differed significantly from the others (as determined by Fisher's LSD post-hoc test, $p < .05$). The longer it took for the change to happen, the longer it also took male participants to detect it.

Taken together, these two analyses are likely suggesting that gender matching may account for some of the variability in our data: women attend more to objects matching their gender, and men do not. However, while the absence of an object gender effect in the male sample might suggest a motivational hypothesis (e.g. familiarity with female objects by a mostly female sample), it is important to note that the male sample consists only of 16 participants.

This analysis may also suggest that motivational differences, if occurring, may focus women's and men's attention on the object, and not on the person in the picture. In both analysis we find no impact of stereotype – neither a direct impact, nor one moderated by threshold.

Perceptive hypothesis

Aims to understand if specific features or details in our materials can lead to observed differences – either promoting effects or interfering with the test of our main hypothesis.

To accomplish this, we first analyzed how participants responded to each specific scene (scenes were counterbalanced between all three levels of threshold). With this goal, average RTs were calculated for each of the 12 scenes displayed; subsequent ANOVAs were performed to detect any interactions between scene and object gender, and scene and stereotype.

Attesting to the relevance of the specific photo, main effects were detected for scene (F (11, 737) = 31.344, $p = .000$, $\eta p^2 = .319$), suggesting that participants took more time to detect changes in specific scenes.

Additionally, even when analyzing by scene, a main effect of object is detected (F (1, 71) = 8.168, p = .006, ηp^2 = .103), but no main effect of stereotype was found (F (1, 67) = 0.005, p < .946, ηp^2 < .001).

Two interactions with scene were detected and plotted: object gender and scene (F (11, 781) = 4.34, p < .001, ηp^2 = .058), and stereotype and scene (F (11, 737) = 3.203, p < .001, ηp^2 = .046).

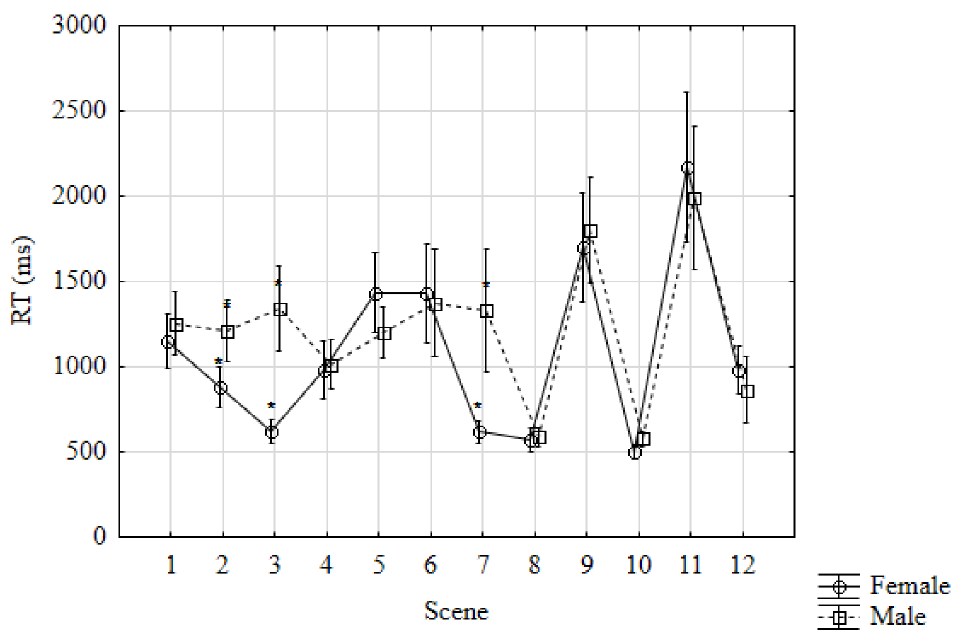


Figure 9. Participants RTs for Object gender by Scene. Asterisks indicate significant differences (Fisher's LSD post-hoc test, with p < .05).

This suggests that only some scenes are responsible for differences found for object gender and stereotype. The plots allow us to single out which scenes had significantly different RTs for these factors, so that their perceptive features, namely saliency differences, could be analyzed.

From post-hoc analysis, we have identified scenes 2, 3, and 7 as showing significantly faster detection of objects associated with women (Figure 9); on the other hand, scene 3 shows faster detection for trials in which a woman is shown, and scene 11 for when a man is shown (Figure 10).

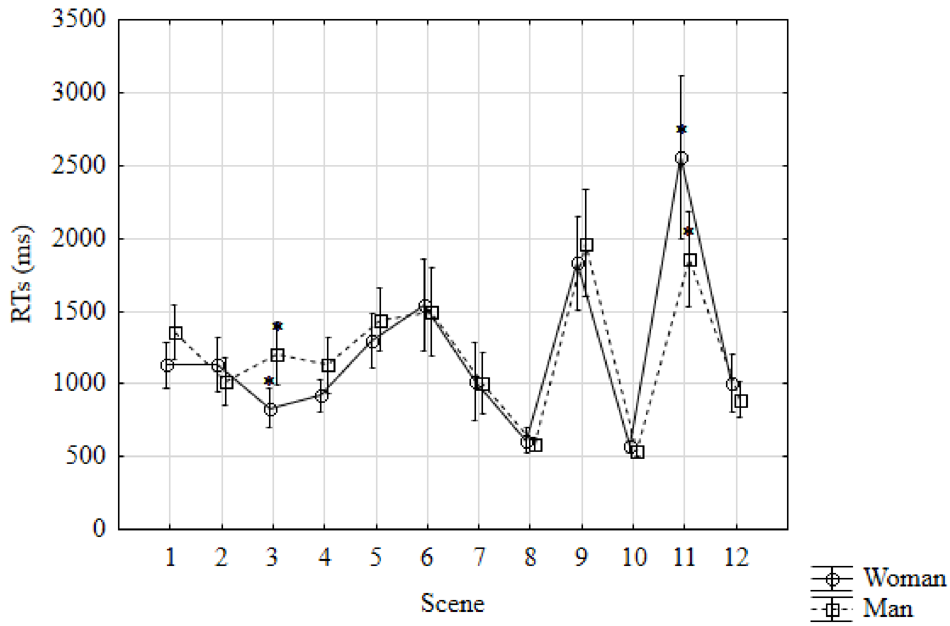


Figure 10. Participants RTs for Stereotype by Scene. Asterisks indicate significant differences (Fisher's LSD post-hoc test, with p < .05).

Saliency analysis

In order to understand if object gender and/or person in the set of photos have different perceptive features that could promote differences in these factors, we analyzed the salience of the images from scenes 2, 3, 7, and 11 – in total, 16 images (4 scenes x 2 object gender x 2 stereotype) – using the SUN model (Zhang, Tong, Marks, Shan, & Cottrell, 2008).

Saliency, to define it briefly (for a more thorough definition, see Appendix A), is that which makes a stimulus attract our gaze or fixation, standing out from other stimuli with which it competes for our attention. It is not an exclusive characteristic of the environment itself, instead depending simultaneously on the observer. In short, something is salient when it implies "biological significance" (Harris & Jenkin, 2001, p. 8), and stimuli characteristics (e.g., orientation and color) can play a part in setting it apart from its context.

Because of this, differences in saliency are relevant in the sense that object (female vs. male) and person (woman vs. man) were changed. These might have introduced new saliency values, altering how salient certain areas of the image are in relation to the rest of the image (i.e., how they compete for our gaze and attention).

The SUN model outputs saliency maps – images in which each pixel is colored according to predicted bottom-up saliency, in relation to the rest of the image. As it uses MATLAB's *imagesc* function, each saliency map is normalized: the most salient pixel in the image is always shown with the brightest yellow; similarly, the darkest blue is always attributed to the pixel scoring lowest in saliency. If pixels with new lower or new higher saliency scores are introduced (i.e., increasing saliency amplitude), the rest of the pixels will be colored according to that new amplitude.

Analysis is, from here onwards, qualitative, by interpreting the saliency maps on the following page (Figure 11).

Object gender

A significant difference for object gender was detected in scenes 2, 3, and 7. Saliency maps for these scenes (Figure 11) show that changing the object can influence the overall saliency of the background, making certain regions of the image more, or less, salient. This could, in turn, make the object easier or harder to detect.

For instance, in scene 3, two aspects can be noted:

a) the female object (left side) has higher saliency in relation to the background, while the male object (on the right) shows levels of saliency (in yellow) that are similar in intensity to the background (i.e., unlike the female object, both male object and its surrounding background show similar color);
b) the introduction of the female object lowers brightness on other areas of the image (e.g., bedframe or top of wardrobe) – this indicates that the female object becomes the most salient aspect of the image, with saliency scores that fall outside the range of saliencies (i.e., higher) for the image with the male object.

In short, the female object is, in relation to the rest of its image, more salient than the male object is. This difference in saliency could render it easier to detect.

In scenes 2 and 7, there is no visible difference in the saliency of the background when the object is changed. Differences in RTs for object gender may result from different saliency scores for object themselves, and the area surrounding them. Saliency scores are calculated by averaging each pixel's individual score – which, in turn, is the algorithm's prediction of how it will attract "human fixation when free-viewing images" (Zhang et al., 2008, p. 2).

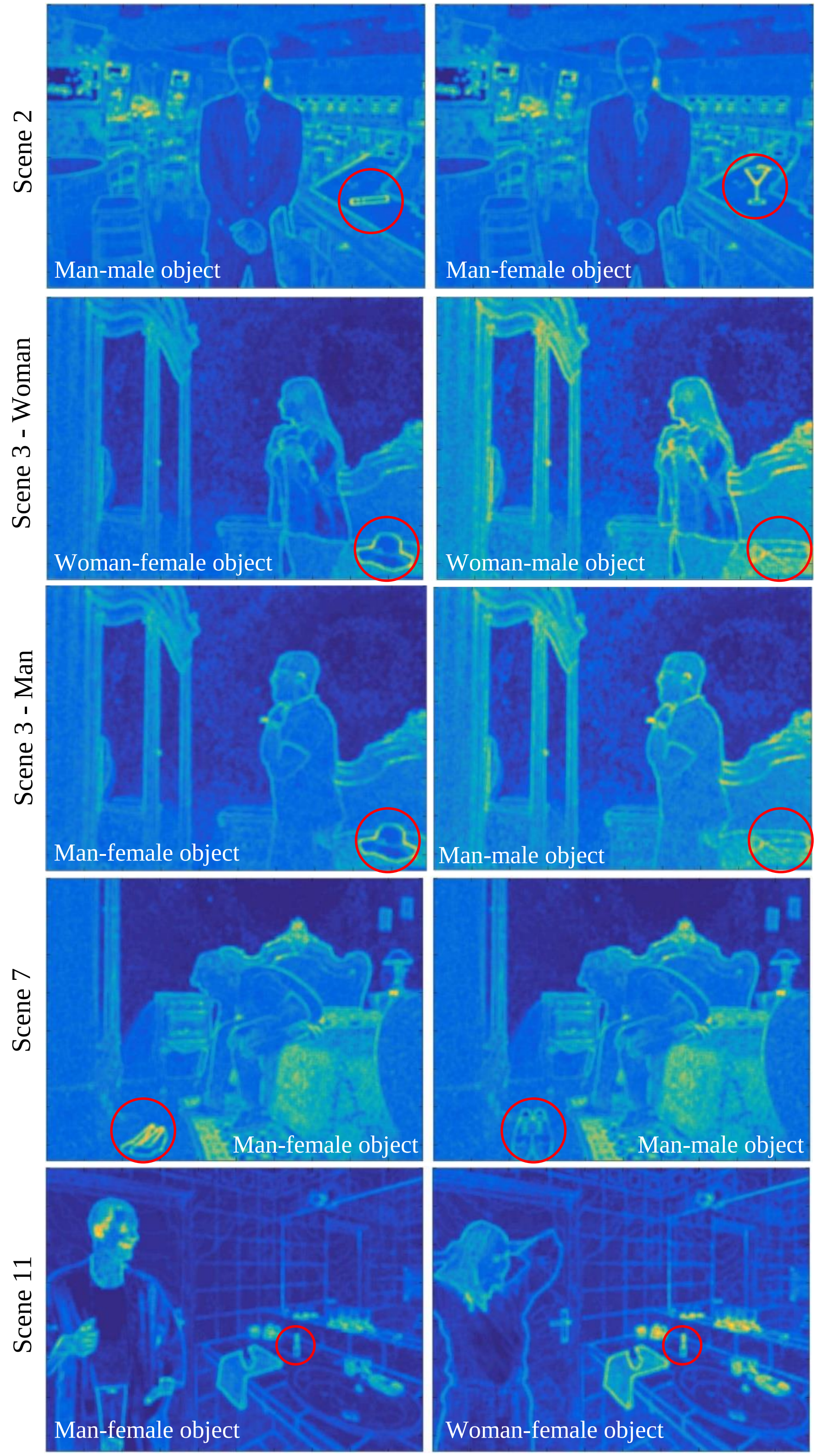


Figure 11. Saliency maps for scenes 2, 3, 7, and 11. Red circles mark object location

In scene 2, for example, these scores were calculated for the region containing the object and its immediate surrounding. The male object score was 160.002 while the female object scored 184.57. This higher salience of the female object could account for its faster detection, regardless of stereotypic context.

As for scene 7, the female object (women's shoes, on the left) is visibly more salient than the male object, as evidenced by its brighter yellow color. In other words, the female object is higher in the ranking of saliency in its own image than the male object is (the blue shades of the men's shoes' pixels match the background's color more closely, meaning it stands out less). This can lead to facilitated detections of the female object.

This analysis tells us, then, that there are aspects of the materials that promote differences in object gender, towards a faster detection of objects associated with the female gender.

Stereotype

As seen above, the different objects in scene 3 also alter the salience of other areas of the image – namely, the person. For both *man* and *woman* conditions, the area of the image containing the person is more salient when the male object is present. However, this happens differently: looking at the male object only, the woman's pixels have brighter yellow tones when compared to the man's – in order words, predicted by the algorithm to attract fixation more often.

In scene 11 the difference in salience is noticeable between man and women, instead of the object itself: the man's face is brighter than the woman's. In fact, brightness for objects on the right side of each image differ: objects are brighter in the *woman* condition – both the critical object and the surrounding objects. On the other hand, in the *man* condition, the face introduces new higher values of saliency, and the objects (including the changing object) are no longer the most salient aspect. In other words, for trials with this scene where a man is featured, the saliency ranking of the object and its surrounding is decreased when compared to trials with a woman. It is difficult to assume, from this analysis of saliency maps, the reason why the *man* condition registered faster detections. One possibility is that the in the *woman* condition both critical and surrounding objects compete for human fixation (i.e., increased noise) – this added information may require increased visual processing and slow down detection.

As demonstrated, scenes 2, 3, 7, and 11 have unique perceptive features that match significant differences in RTs for either object gender or stereotype. To rule out the possibility that these images may have introduced noise in our main analysis, trials from scenes 2, 3, 7, and 11 were removed and the analysis of our hypothesis was repeated.

Re-testing the hypothesis controlling for saliency differences

To investigate the impact of the materials and any perceptive features that could interfere with participants' performance, scene saliency was analyzed. Scenes 2, 3, 7, and 11 had significant differences in RTs between either object gender, or stereotype. Analysis of saliency maps for these scenes also showed notable differences between conditions (e.g., female object vs. male object). To rule out the perceptive hypothesis for the object gender effect, trials from these four scenes were removed, and our main hypothesis analysis was repeated.

We analyzed participant RTs for each stereotype, object gender, and threshold, as within-participant factors of an ANOVA with the following design: 2 (woman vs. man) x 2 (female vs. male) x 3 (400ms, 800ms, 1200ms).

Only an effect of threshold was detected (F (2,138) = 37.663, p = .000, ηp^2 = .353). No other effects were observed. Interestingly, by controlling for these scenes, we find that RTs at 1200ms were significantly higher than at 400ms and 800ms, as determined by Fisher-s LSD post-hoc test (p = .000).

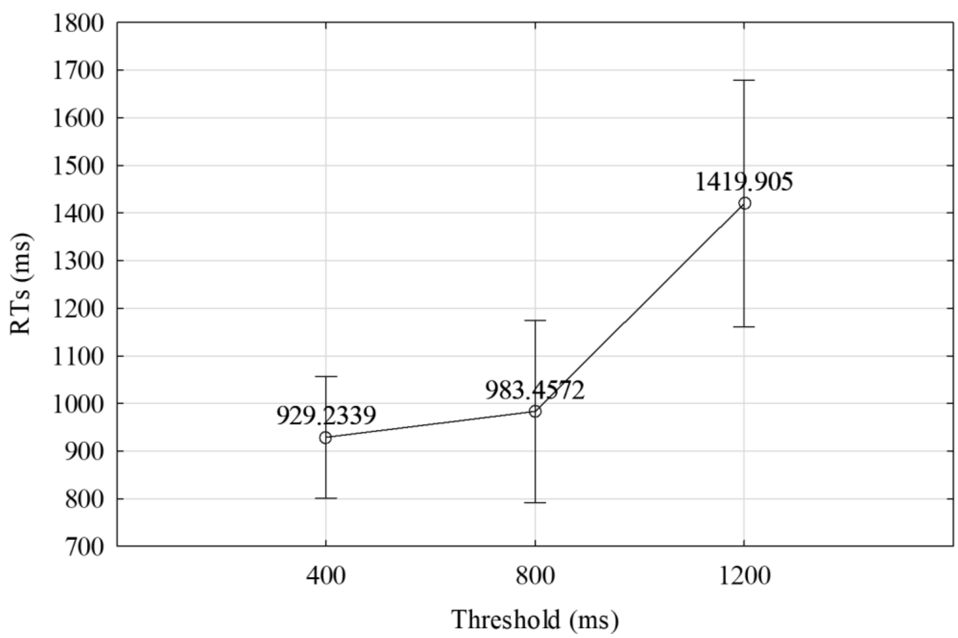


Figure 12. RTs for all levels of threshold, after removing scenes 2, 3, 7, and 11.

Analysis of threshold effects

The only consistent effect in our data is that of threshold. What can explain this?

It should be noted that in the version of the flicker paradigm used here, both images (whether with or without the changing object) as well as the blank screens were displayed for 200ms. As such, it is possible to determine where in the procedure participants detected the change (i.e., when an image was shown with the object, without the object, or during a blank screen). Adding the average RT and its corresponding threshold gives us the total time from the start of the procedure. Figure 16 is a graphical representation of the procedure's timeline.

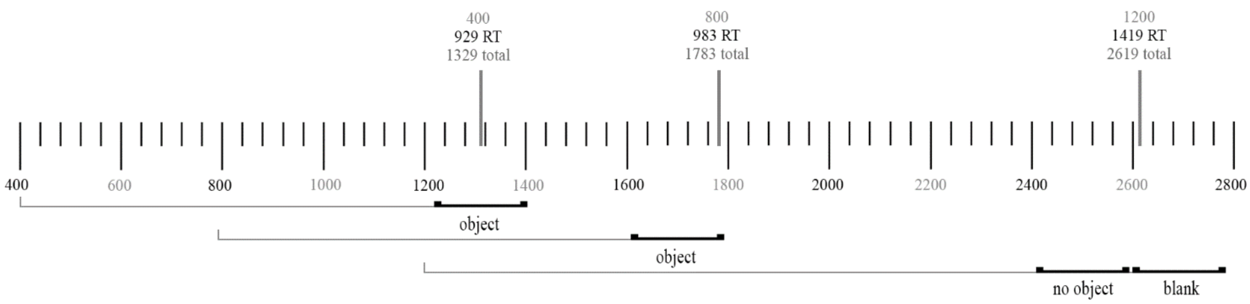


Figure 13. Timeline (in milliseconds) of the procedure, indicating average RT for each threshold condition.

The top layer shows average RTs registered for each threshold level (identified in gray, top line). Total RTs are the sum of the observed RT and the threshold level (e.g. 1329 = 929 + 400). In the bottom part, each red line is associated with a specific threshold, and indicates whether detection happened in a blank screen, or an image with or without the changing object. The ruler in the middle section distinguishes between an image being presented (number in black), or a blank screen (number in gray).

As such, average RT for threshold levels of 400ms and 800ms land on a picture containing the changing object (i.e., the object "appears"); more specifically, average RT for 400ms trials is 129ms after the object appears, and for 800ms trials is 183ms after the object appears. Average RT for a threshold level of 1200ms lands on a blank screen that was preceded by an image without the changing object; in other words, at 1200ms trials, the key is pressed 219ms after the object disappears.

To try to understand if the threshold effect was limited to our study and task (e.g., due to error), we collected the threshold values used in experiments from other studies, as well as the RTs observed. Included are studies that used the flicker paradigm and reported RTs in milliseconds (some reported screen cycles, or only show RTs graphically). Some studies reported RTs by condition (e.g., congruency with background), instead of generally, and that distinction was maintained.

Table 1

Studies using the flicker paradigm, thresholds, and RTs.

Study	Experiment	Threshold (ms)	RTs (ms)
Hollingworth and Henderson (2000)	1	330	1261
		330	1190
Zimmermann, Schnier, and Lappe (2010)	1	320	4500
LaPointe et al. (2013)	3	500	5126
LaPointe and Milliken (2016)	–	500	4778
		500	3761
Ortiz-Tudela, Milliken, Botta, LaPointe, and Lupiañez (2016)	1	500	3289
		500	2680
	2	500	3072
		500	2407
	3	500	4018
		500	3691
LaPointe and Milliken (2017)	1	500	4368
		500	3638
Vierck and Kiesel (2008)	1	320	8156
		640	8917
		1600	15504
	2	960	11746
		1600	16020
	3	640	8583
Hollingworth, Schrock, and Henderson (2001)	1	660	2760
	2a	660	2930
	2b	660	5050

There is a significant correlation between thresholds used and the observed RTs (Pearson product-moment correlation; r (21) = .86, p = .000). Note that only the study by Vierck and Kiesel (2008) employs three different thresholds for the same task, with the same stimuli. It is also the only study using relatively high threshold values (namely, 1600ms and 960ms). Removing this study from the analysis renders the correlation insignificant (r (21) = .328, p = .199)

Discussion

This study aimed to test if congruency between the gender associated with a changing object and an exemplar of the gender stereotype in a flashing image would interfere with the speed at which participants detected the change in a visual task. As such, we looked for an interaction effect between stereotype and object gender (congruency). Such effect was not found.

An effect of object gender (faster detection for female objects) was observed in the first stage of our analysis. Exploring genders separately showed it to be present only in the female sample. Scenes in which this effect was significant were identified and their perceptive qualities were analyzed with saliency maps. Due to visible discrepancies in saliency between the two object conditions, these scenes were removed, and the analysis repeated for the full sample. No effect of object gender was detected.

Finally, the only significant effect was that of threshold: change detection is significantly slower when change occurs at 1200ms.

Congruency

Throughout every stage of the analysis, no such interaction was observed in our results, suggesting that stereotype-object congruency had no impact in change detection performance. A few possible explanations can be advanced.

Reviewing available literature on visual search tasks and the effect of context and expectations, particularly when tested in conjunction with stimulus-context congruency or including stereotypes, some observations stand out:

a) For *identification* tasks, a congruency benefit is observed, both with neutral stimuli (e.g. a road, university campus, or a zoo; LaPointe, Lupianez, & Milliken, 2013; LaPointe & Milliken, 2016, 2017; Ortiz-Tudela, Milliken, Botta, LaPointe, & Lupiañez, 2016), and using stereotypes as context (Correll et al., 2002; Payne, 2001) – in short, stereotypes acted as other schemas that are abstracted from the image's background;
b) For *detection* tasks, the benefit (i.e., faster detection) is observed for neutral incongruent stimuli (experiments 2 and 3 by LaPointe et al., 2013; Ortiz-Tudela et al., 2016).

> However, no evidence was found on how stereotyped contexts would interact with congruency. This gap in the literature is why we approached the issue.

Expecting an interaction (as we hypothesized) contains an assumption that has not been verified: that if stereotypes behave similarly to other (neutral) schemas in identification tasks, they would also perform similarly for detection tasks. There is a possibility that task type (i.e., identification vs. detection) might trigger cognitive mechanisms that impact the processing of stereotype-heavy stimuli differently. Stereotypes could operate differently than other schemas in detection tasks, while operating similarly in identification tasks. But even when limiting our focus to identification tasks, it is not necessary that, cognitively, stereotypes act as other schemas to obtain the same result – a different mechanism, acting with stereotypes but not neutral stimuli, could also lead to a congruency benefit. In short, similar results do not guarantee similar processes.

LaPointe and Milliken (2017), when detailing their model, attempt to explain both the congruency benefit in identification task, and the incongruency benefit in detection tasks (see Figure 1 in the Introduction section). To do so, they suggest different mechanisms for the two tasks: identification relies, for example, on prior knowledge and expertise; while detection relies on *attention capture processes* that focus on semantic inconsistency. It is possible to argue that stereotypes rely differently on such mechanisms, with heavier dependency on prior knowledge – as said by Hilton and Hippel (1996, p. 248) in their text on stereotypes, "prior knowledge determines what we see and hear, how we interpret that information, and how we store it for later use". This would explain how stereotypes seem to make a difference in identification tasks, but not in detection tasks: stereotypical incongruency might not interact with the attention capture processes with the strength that translates to a differential reaction time.

A second problem stems from differences in method. While studies approaching the impact of congruency have adopted the flicker paradigm (also used in this study; Rensink et al. 1997), studies involving stereotypes have not. In the case of Payne's study (2001), the task was more a traditional priming paradigm than a visual search task: faces were shown in isolation, without any background or attempt to make them naturalistic. It is legitimate to ask if different methods might be demonstrating, once again, different cognitive processing that just happen to output the same result – congruency benefits. As such, the question that follows is: if the same method was applied (e.g., the flicker paradigm) to identification tasks with stereotypes, would the same benefit be observed? In other words: under the same method, would stereotypes act as other neutral schemas do? Note that this was an assumption that sustained our research question.

Under this perspective, our results (no effect of congruency) might indicate that with stereotype-activating contexts, congruency has no effect on performance in a change detection task. However, it might also be indicative that this lack of effect is not dependent on task type – a question to be cleared by running a change identification task based on the flicker paradigm, using stereotypes as context.

Concerning the lack of effect for congruency, an aspect worth noting regards the materials used. Recall that the images presented to participants contained an exemplar of the gender stereotype (a man or a woman) in a naturalistic background, such as a living room or a garden. The changing objects, while varying in congruency with the stereotype exemplar, were carefully selected to maintain congruency with the backgrounds – a control measure to ensure that differences in RTs were due to congruency with the stereotype and not with the backgrounds. However, our study has no means of determining if, either during gist processing or later in the trial, participants were in fact attending (and creating expectations from) the stereotypes or the backgrounds. There is a possibility that the scenery (e.g., bathroom), and not the stereotype, took over the role of context – in this situation, congruency between object and stereotype would not be processed, as it would be overridden by congruency between object and scenery. All trials would, then, be congruent trials, and no congruency or incongruency effect involving object would be detected.

Choice of stereotype must also be discussed. Gender is typically associated with specific roles (in fact, commonly known as *gender roles*), in which are included the different behaviors or characteristics expected of each gender, that are passed on from generation to generation (Neculăesei, 2015). As such, expectations regarding gender vary from culture to culture (e.g., man are leaders, women are maternal). Hofstede (2017) explores, among other characteristics, how different societies differ regarding how strict such gender roles are: those in which the roles are polarized and distinct are labeled Masculine; those in which the roles overlap are named Feminine. On this axis, according to the author, Portugal has a low score of 31 – a more feminine culture. With this in mind, a lack of a congruency effect might result from the fact that expectations regarding gender roles are not strong or exclusive enough to trigger a distinct response to congruent and incongruent trials. For example, a man next to women's shoes might be recognized, upon deliberation, as somewhat inconsistent (after all, the objects were strongly associated with one gender in pre-test), but still not strike the participant as something out of the ordinary, or deserving more attention and urgency.

Object gender

The next initial finding was an effect of object gender, in which female objects were detected significantly faster than male objects. A hypothesis based on familiarity was proposed: considering that the sample was largely female (78%), could female objects be easier to recognize and detect? In short, can familiarity (or, considering we used everyday objects, *superfamiliarity;* Buttle & Raymond, 2003) improve detection? This motivational hypothesis is not a novel factor, with other studies already suggesting that expertise (which implies familiarity) can inhibit change blindness under the flicker paradigm (Curran, Gibson, Horne, Young, & Bozell, 2009; Jones, Jones, Smith, & Copley, 2003; Werner & Thies, 2000), resulting in lower RTs for changes using familiar objects and scenes. To explore this possibility further, we conducted separate analyses for our female and male participants and observed that the benefit for female objects was found in the female sample, but not in the male sample. It is important to highlight, however, that our male sample had only 16 male participants – the results cannot be taken as conclusive.

The object gender effect can also be linked to the materials – namely, we hypothesized that differences in object saliency might attract eye gaze differently (Boyer, Smith, Yu, & Bertenthal, 2011; Underwood & Foulsham, 2006; Underwood, Foulsham, van Loon, Humphreys, & Bloyce, 2006; Wright, 2005). To test this perceptive hypothesis, we focused on scenes that registered significantly different RTs for male and female objects – in all, the same benefit for female objects was observed. Saliency maps and scores created for these scenes provided a viable alternative explanation for the observed faster detection of female objects. After removal of these scenes, no effect of object gender was found, strengthening the perceptive hypothesis.

Another reason why the perceptive hypothesis prevails over the familiarity hypothesis is that participants were instructed to merely *detect*, even when unable to name the change (as is typical of identifications tasks). Unlike a correct identification, a correct detection can occur even when the object is not fully registered in subjective awareness but instead only sensed (Rensink, 2004). Because the current study provides no data regarding eye gaze, detections without directly looking at the changing object are a possibility that cannot be discarded. Without solid data on how often participants directed their gaze at the changing objects, effectively perceiving them, a hypothesis based on familiarity cannot be safely advanced. Considering evidence for implicit memory (see Schacter, 1987), it can be argued that recognition and familiarity can occur even outside of conscious awareness of detailed visual perception. It can even be argued that, having used everyday objects, previous exposure had been established. But if the effect of object gender could be attributed to familiarity, we would expect it to be expressed significantly in all the scenes, instead of only in three (25%) that also had differing saliency configurations.

Threshold

The only effect consistently detected in every stage of the analysis was that of threshold. After removal of the four scenes with saliency issues, change occurring at 1200ms was detected significantly slower when compared to the two other conditions (400ms and 800ms). The three different thresholds were created to detect and eliminate participants who did not follow instructions, and as such was an original addition to the procedure and the flicker task. Our literature review found only one study with a similar manipulation of threshold, as defined by the time between the start of a trial, and the moment when change first occurs. Focusing on whether visual information is integrated along repeated exposures into a long-term memory representations, Vierck and Kiesel (2008) recorded RTs in a variety of configurations of original (A) and altered images (B). Experiment 1 compared conditions AB vs. AABB vs. AAAAABBBBB – much like the present study, the change happens increasingly later, with more displays of the original unmodified image filling in for the time until change occurs. Also similarly to the present study, the highest threshold (AAAAABBBBB condition) registered significantly higher RTs, with the two first conditions registering no significant difference between them. Note, however, that the alternation does not return to an ABAB format once change happens; instead, the authors purposely defined each cycle to contain as many repetitions of A as of B. From this method and the observed RTs, they conclude: "repetitions are more beneficial for change detection when they affect both the original and the modified versions of the stimulus material." (Vierck & Kiesel, 2008, p. 320) Experiment 3 compares an AAABBB cycle with an AAAAAB cycle – this last one being similar to the procedure used in our study. Once more, the condition with more repetitions of the original image yielded higher RTs (i.e., worse performance and slower detection). The authors interpret these results as evidence of "the importance of presenting both picture versions for a prolonged time” (p. 321) to allow for comparable representations in long-term memory, linking change detection performance to those representations. Our results do not contradict this interpretation – our 400, 800, and 1200ms conditions can be translated, respectively, to AB, AAB, and AAAB. As the imbalance between the two versions increased, so did RTs.

However, there may be an alternative to the “long exposure of both images” explanation that accounts for both Vierck and Kiesel's (2008) and our own results. If repetition of the unchanged image (A) continuously adds information to memory, this progressive complexifying of our pre-change visual representation might be more cognitively taxing, requiring more resources, which would be translated to longer RTs. For instance, retrieval of a specific area of an image (to compare pre- and post-change) might be made more difficult due to the quantitative increase of stimulus to be attended to and encoded during longer thresholds. It is known, in memory literature, that increasing a list of items in a memory set can lead to longer retrieval times (Sternberg, 1969); as for visual memory and change blindness, limiting the amounts of items to be retrieved for comparison, usually by cueing something about the upcoming change (such as its location), seems to improve detection performance (Hollingworth, 2003). In this case, repeating the pre-change version of the image might just be adding to the list of objects or features to be retrieved for comparison, slowing down the detection process.

A different proposal can be advanced to account for our results regarding threshold: a change in strategy. We have outlined the importance of gist processing, how it is triggered early, occurring in the first 150ms after presentation of an image (Biederman, 1981; Sampanes, Tseng, & Bridgeman, 2008) and guiding our search for objects according to expectations. However, it is possible that failing to detect change, or to find a target object, the gist strategy is eventually replaced by different strategies. If the amount of time allocated to gist processing is shorter than that during which the unedited image is repeated (e.g. within the first 1200ms of the presentation of the first image), change could occur when a different strategy is already in place. Tatler, Gilchrist, and Rusted (2003) attempted to construct a timeline of what we attend to, and how we build visual representations – what happens when we are exposed to a scene, and in what order? Subjects in their study were shown an image for durations ranging between 1 and 10 seconds. Afterwards, they were asked questions about the image, covering information regarding: gist, presence, shape, color, position, and relative distance. By crossing information from the duration of the images with performance on the follow-up questions, they were able to conclude that:

a) Performance for gist-related questions peaks at 1s, remaining unchanged with longer exposure. At 1s, performance for all other questions is still at chance level;
b) At 2s, performance in questions regarding absolute position is above chance; all other categories (e.g., presence) seem to require longer exposures;
c) For presence and shape information, performance "has not reached an obvious plateau even after 10s" (p. 589).

In short, gist processing seems to be the first strategy, triggered as soon as 150ms (Biederman, 1981) and keeping active up to first second; information about absolute position is attended to at 2s of exposure; information regarding presence of objects in the image can take as long as 10s to be adequately attended to, in order to translate to above-chance performance in follow-up questions. Note that Tatler et al. (2003) is probing explicit memory. As we have reviewed here (see Appendix A), detection is not necessarily explicit or able to be reported on explicitly. In fact, we have reviewed evidence of the impact of expectations, and congruency effects with neutral stimuli, many being registered at times much shorter than 2s or 10s. It is hard to determine if the times reported by Tatler et al. (2003) put the theory of a timeline of different strategies into question, or if they reflect limitations in recalling information instead.

Nevertheless, the debate over global vs. local processing is not new. Seminal work by Navon (1977, p. 354) argues that global processing takes temporal precedence, "in a process of being zoomed in on, where at first it is relatively indistinct and then it gets clearer and sharper." In review, Kimchi (1992) noted that many factors can influence how perception occurs, but concludes that, all things being equal, global processing usually precedes local processing, with evidence that it occurs early in the moment of perception. In short, research on responses to visual stimuli seem to converge on a two-phase model (Gale, 1997, p. 131): a first and early stage, characterized by "preattentive or distributed" visual processing; and a subsequent stage, "focal attention", processing details from a specific area, and display scanning with eye movement.

From this, we propose that a change of strategy, from global (gist) to local, can account for the slower detection in our 1200ms condition. Gist processing, at work during the 400ms and 800ms marks, attends to the globality of the image, dispersing attention in a way that increases the probability of detecting a changing object. Local processing, however, focuses attention on specific details or regions of the image; for the changing object to be detected, change would have to occur in the area being observed, effectively lowering the probability of detecting change early in the process. This proposal can be tested by replicating the study and adding eye-tracking measures, making it possible to identify the dispersion of attention typical of gist-processing, as well as the focalized attention that would follow it.

Finally, a change of strategy might combine with how we encode and compare information across scenes to explain our threshold effect. Hollingworth and Henderson (2002) report that changes to objects that were previously fixated were detected at a higher rate, (vs. changes to non-fixated objects). Regarding how this may impact a flicker task, Hollingworth (2006, p. 794) suggests that several seconds may be needed before each potential changing object is fixated on, "explaining delays in detection of repeating changes, such as those in the flicker paradigm".

But why would this affect longer thresholds (1200ms) and not shorter ones (400ms and 800ms)? One possibility is that the switch from gist processing to a localized, focalized search also lowers the likelihood of change occurring in a previously fixated area (hence improving detection) – after all, in local search, attention moves from object to object, or region to region, and only by chance would a relevant region (i.e., where change occurs) be fixated before all others in the image.

Limitations and further research

Our findings contribute to filling a gap in knowledge regarding how stereotypes-as-context impact congruent and incongruent changes in a change blindness task. Congruency effects previously demonstrated in flicker tasks using neutral stimuli were not observed with our stereotyped backgrounds. This opens the possibility that, while being schemas, stereotypes act differently on congruency according to the type of task. While shoot/don't shoot tasks seem to register a clear impact of stereotype-object congruency (Correll et al., 2002), our task did not – in fact, stereotype alone did not register a significant effect either. As that studies in change detection using stereotypes and congruency are either scarce or nonexistent, our results alone cannot be read as robust evidence that stereotype-object congruency has no impact. However, three possibilities can be raised from our study: stereotype-object matching does not interfere with the specific phenomenon of change blindness; the flicker paradigm is inadequate to detect it; the choice of gender stereotype is inadequate, either for activating other interfering processes or by excessive overlap in gender roles.

Further research should tackle the open question regarding a possible effect if we had used an *identification* version of this task (since the studies we reviewed that used stereotypes asked participants to identify critical stimuli). This could be settled by changing the instructions shown to participants and by adding a selection screen between each trial to collect responses regarding the identity of the changing objects.

The possibility that congruency between object and setting may override congruency (or incongruency) with the stereotype should also be addressed in future research. Using naturalistic stimuli required depicting familiar locations, such as a bathroom or a garden. These, consequently, activate a series of representations of objects that we usually expect to see in such settings. Clarification is needed in terms of what is more dominant when we observe images such as the ones used here: processing of congruency with the setting, or processing of congruency with an exemplar of a stereotype? The literature here reviewed tells us that a photograph of a bedroom can activate the representation of clothing, and even influence the speed at which it is detected. But adding an exemplar of a stereotype complicates matters: if a man is seen in the bedroom, men's shoes would be congruent with both setting and stereotype – but women's shoes would be congruent with the setting, while incongruent with the stereotype. Which one is primarily processed? The effect of stereotype-object congruency becomes harder to isolate. As such, further explorations of this topic should take this issue into account and include conditions of incongruency with the background. Ideally, a task could be setup to establish which of the congruencies is dominant – if any – but it would require backgrounds from which expectations are impossible, or at least unlikely. In other words, a step away from naturalistic stimuli.

Limitations regarding method can also be addressed. Saliency, for instance, was considered only after data collection. Analyzing saliency maps beforehand would allow for more precise control over that confounding factor. The same control could be considered for objects' shapes and colors.

Some questions raised by our results could also be answered by adding an eye-tracking component to the experiment. Gaze information would allow for precise information regarding where participants attended (e.g., did they attend to stereotypes or random objects in the setting?); more importantly, it would yield information regarding the question of gist vs. local visual search, possibly providing explanatory evidence of the significant differences in RT observed between threshold conditions. Naturally, these limitations can also be read as details to consider (and valuable additions) to future research.

References

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (1st ed., pp. 213–233). London: Routledge.

Boyer, T. W., Smith, T. G., Yu, C., & Bertenthal, B. I. (2011). Visual attention and change detection. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1735–1740).

Bracco, F., & Chiorri, C. (2009). People have the power: Priority of socially relevant stimuli in a change detection task. *Cognitive Processing*, *10*(1), 41–49. https://doi.org/10.1007/s10339-008-0246-7

Busch, N. A., Fründ, I., & Herrmann, C. S. (2010). Electrophysiological evidence for different types of change detection and change blindness. *Journal of Cognitive Neuroscience*, *22*(8), 1852–1869. https://doi.org/10.1162/jocn.2009.21294

Buttle, H., & Raymond, J. E. (2003). High familiarity enhances visual change detection for face stimuli. *Perception & Psychophysics*, *65*(8), 1296–1306. https://doi.org/10.3758/BF03194853

Cadinu, M., Maas, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology*, *33*, 267–285. https://doi.org/10.1002/ejsp.145

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, *4*(5), 170–178. https://doi.org/10.1016/S1364-6613(00)01476-5

Correll, J., Park, B., Judd, C., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, *83*(6), 1314–1329. https://doi.org/10.1037/0022-3514.83.6.1314

Cox, W. T. L., Abramson, L. Y., Devine, P. G., & Hollon, S. D. (2012). Stereotypes, prejudice, and depression: The integrated perspective. *Perspectives on Psychological Science*, *7*(5), 427–449. https://doi.org/10.1177/1745691612455204

Curran, T., Gibson, L., Horne, J. H., Young, B., & Bozell, A. P. (2009). Expert image analysts show enhanced visual processing in change detection. *Psychonomic Bulletin and Review*, *16*(2), 390–397. https://doi.org/10.3758/PBR.16.2.390

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing black: Race, crime, and visual processing. *Journal of Personality and Social Psychology*, *87*(6), 876–893. https://doi.org/10.1037/0022-3514.87.6.876

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. https://doi.org/10.3758/BF03193146

Gale, A. G. (1997). Human response to visual stimuli. In W. R. Hendee & P. N. T. Wells (Eds.), *The perception of visual information* (pp. 127–147). New York, NY: Springer.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4–27. https://doi.org/10.1037/0033-295X.102.1.4

Harris, L. R., & Jenkin, M. (2001). Vision and attention. In M. Jenkin & L. R. Harris (Eds.), *Vision and attention* (pp. 1–17). New York: Springer.

Hilton, J. L., & Hippel, W. Von. (1996). Stereotypes. *Annual Review of Psychology*, *47*, 237–271.

Hofstede, G. (2017). Geert Hofstede - Geert Hofstede. Retrieved from https://geert-hofstede.com/

Hollingworth, A. (2003). Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 388–403. https://doi.org/10.1037/0096-1523.29.2.388

Hollingworth, A. (2006). Visual memory for natural scenes: Evidence from change detection and visual search. *Visual Cognition*, *14*(4–8), 781–807. https://doi.org/10.1080/13506280500193818

Hollingworth, A., & Henderson, J. M. (2000). Semantic informativeness mediates the detection of changes in natural scenes. *Visual Cognition*, *7*(1–3), 213–235. https://doi.org/10.1080/135062800394775

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(1), 113–136. https://doi.org/10.1037/0096-1523.28.1.113

Hollingworth, A., Schrock, G., & Henderson, J. M. (2001). Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory and Cognition*, *29*(2), 296–

304. https://doi.org/10.3758/BF03194923

Javadi, A. H., & Wee, N. (2012). Cross-category adaptation: Objects produce gender adaptation in the perception of faces. *PLoS ONE*, *7*(9), 3–10. https://doi.org/10.1371/journal.pone.0046079

Jones, B. T., Jones, B. C., Smith, H., & Copley, N. (2003). A flicker paradigm for inducing change blindness reveals alcohol and cannabis information processing biases in social users, 235–244.

Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin*, *112*(1), 24–38.

LaPointe, M. R. P., Lupianez, J., & Milliken, B. (2013). Context congruency effects in change detection: Opposing effects on detection and identification. *Visual Cognition*, *21*(1), 99–122. https://doi.org/10.1080/13506285.2013.787133

LaPointe, M. R. P., & Milliken, B. (2016). Semantically incongruent objects attract eye gaze when viewing scenes for change. *Visual Cognition*, *24*(1), 1–15. https://doi.org/10.1080/13506285.2016.1185070

LaPointe, M. R. P., & Milliken, B. (2017). Conflicting effects of context in change detection and visual search: A dual process account. *Canadian Journal of Experimental Psychology, 71*(1), 40–51. https://doi.org/10.1037/cep0000105

Lyyra, P., Wikgren, J., & Astikainen, P. (2010). Event-related potentials reveal rapid registration of features of infrequent changes during change blindness. *Behavioral and Brain Functions*, *6*(12), 1–7. https://doi.org/10.1186/1744-9081-6-12

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology, 66*(1), 37–47. https://doi.org/10.1037/0022-3514.66.1.37

McCallum, R., & Lucas, G. (1999). *Star Wars. Episode I, The Phantom Menace*. United States: 20th Century Fox.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–383.

Neculăesei, A.-N. (2015). Culture and gender role differences. *Cross-Cultural Management Journal*, *17*(1), 31–35.

Ortiz-Tudela, J., Milliken, B., Botta, F., LaPointe, M., & Lupiañez, J. (2017). A cow on the prairie vs. a cow on the street: Long-term consequences of semantic conflict on episodic encoding. *Psychological Research*, *81*(6), 1264–1275. https://doi.org/10.1007/s00426-016-0805-y

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*(2), 181–192. https://doi.org/10.1037/0022-3514.81.2.181

Rensink, R. A. (2002). Changes. *Progress in Brain Research*, *140*, 197–207. https://doi.org/10.1016/S0079-6123(02)40051-9

Rensink, R. A. (2004). Visual sensing without seeing. *Psychological Science*, *15*(1), 27–32. https://doi.org/10.1111/j.0963-7214.2004.01501005.x

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368–373. https://doi.org/10.1111/j.1467-9280.1997.tb00427.x

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*(4), 331–363. https://doi.org/10.1037/1089-2680.7.4.331

Sampanes, A. C., Tseng, P., & Bridgeman, B. (2008). The role of gist in scene recognition. *Vision Research*, *48*(21), 2275–2283. https://doi.org/10.1016/j.visres.2008.07.011

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *13*(3), 501–518. https://doi.org/10.1037//0278-7393.13.3.501

Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition*, *7*(1–3), 1–15. https://doi.org/10.1080/135062800394658

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inatteintional blindness for dynamic events. *Perception*, *28*, 1059–1074. https://doi.org/10.1068/p2952

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*(7), 261–267.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, *5*(4), 644–649.

https://doi.org/10.3758/BF03208840

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9*(1), 16–20. https://doi.org/10.1016/j.tics.2004.11.006

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57*(4), 421–457.

Stone, J., Perry, Z. W., & Darley, J. M. (1997). "White men can't jump": Evidence for the perceptual confirmation of racial stereotypes following a basketball game. *Basic and Applied Social Psychology*, *19*(3), 291–306. https://doi.org/10.1207/15324839751036977

Straube, S., & Fahle, M. (2011). Visual detection and identification are not the same: Evidence from psychophysics and fMRI. *Brain and Cognition, 75*(1), 29–38. https://doi.org/10.1016/j.bandc.2010.10.004

Tatler, B. W., Gilchrist, I. D., & Rusted, J. (2003). The time course of abstract visual representation. *Perception*, *32*, 579–592. https://doi.org/10.1068/p3396

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology*, *59*(11), 1931–1949. https://doi.org/10.1080/17470210500416342

Underwood, G., Foulsham, T., van Loon, E., Humphreys, L., & Bloyce, J. (2006). Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology*, *18*(3), 321–342. https://doi.org/10.1080/09541440500236661

Vierck, E., & Kiesel, A. (2008). Change detection: Evidence for information accumulation in flicker paradigms. *Acta Psychologica*, *127*(2), 309–323. https://doi.org/10.1016/j.actpsy.2007.06.004

Werner, S., & Thies, B. (2000). Is "change blindness" attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition, 7*(1–3), 163–173. https://doi.org/10.1080/135062800394748

Wright, M. J. (2005). Saliency predicts change detection in pictures of natural scenes. *Spatial Vision*, *18*(4), 413–430. https://doi.org/10.1163/1568568054389633

Zhang, L., Tong, M. H., Marks, T. K., Shan, H., & Cottrell, G. W. (2008). SUN: A Bayesian framework for saliency using natural statistics. *Journal of Vision*, *8*(7), 32.

https://doi.org/10.1167/8.7.32

Zimmermann, E., Schnier, F., & Lappe, M. (2010). The contribution of scene context on change detection performance. *Vision Research, 50*(20), 2062–2068. https://doi.org/10.1016/j.visres.2010.07.019

Appendices

Appendix A – Literature Review

Change blindness - what it is

There is little variance in how change blindness is defined in scientific literature – in short, it is the "the inability to detect changes to an object or scene" (Simons & Levin, 1997, p. 261). While many variations of the phenomenon have been explored in laboratory settings (for a review, see Simons & Ambinder, 2005; Simons & Rensink, 2005), it also happens frequently in our everyday life. One relatable example is known in the movie industry as a continuity error: a lack of consistency in a scene's objects, characters, or details, usually occurring between cuts, and very infrequently noticed (as illustrated in Figure A1).

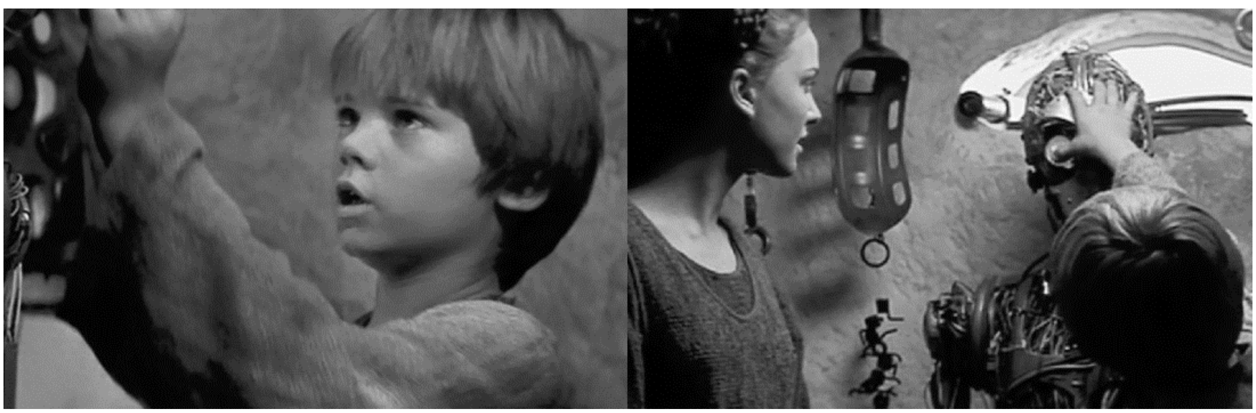


Figure A1. Continuity error in the popular film Star Wars. Episode I, The Phantom Menace. Anakin reaches for the android with his left arm in the first shot, but is seen using his right arm immediately after the camera cuts (McCallum & Lucas, 1999).

The example of continuity errors in movies reflects how pervasive and unnoticed change blindness is. After all, every movie has multiple of such errors, and yet most of the audience barely ever notices one. Even those involved in moviemaking fail to notice such visual inconsistencies, despite being repeatedly exposed to the material. We may think we are good at detecting changes, but instead we appear to be blind to our own change blindness – change blindness blindness (Levin, Drivdahl, Momen, & Beck, 2002; Levin, Momen, Drivdahl, & Simons, 2000).

Defining change blindness implies defining the concept of *change* to an operable level. Moreover, change can, indeed, be distinguished from other types of visual phenomena that can also be missed or detected by an observer. Rensink (2002) underlines the importance of this distinction by noting how separate visual phenomena are often based on different perceptual processes, sometimes correlated, but also able to be detected independently by adequate methods.

When the author uses the word *change*, it is in reference to a "transformation or modification of something over time" (p. 248) and, as such, distinct from *motion* – an also detectable phenomenon, but consisting of a modification of position, or location, by a determined and measurable quantity. With *motion*, there is no direct modification of an entity or structure itself – a central requirement in the definition of *change* within the context of change blindness.

Change can also be contrasted with *difference* in the sense that the former is relative to one single entity, while the latter refers to dissimilarity between more than one entity. Consequently, *difference* can be determined by atemporal comparison of the two separate entities, while *change* implies transformation and, therefore, different moments in time (a before and an after). This temporal requirement allows us to also distinguish between two types of changes that can be observed: one that is in progress, perceived as ongoing, in the present – in short, a *dynamic change* (imagine an object that slowly fades in or out of a picture; e.g. Simons, Franconeri, & Reimer, 2000); and one that has already occurred in the past, unwitnessed in its progress between a previous and a present state, and perceived as a *completed change*.

What exactly happens?

Why is it apparently hard to notice something changing from one visual scene to the next? If we perceive and process both scenes, why are we not always alerted to a difference between the two? Or do we even process both scenes?

In a review of the research surrounding change blindness, Simons (2000) describes standing explanations to the phenomenon – each suggesting different ways in which information from visual scenes is processed. While the author attempts to list "all of the plausible models that have been considered" (p. 8), it is important to underline that no single theory can account for *all* findings in the field, and whatever mechanism is in action may also depend on a study's method, stimuli, and context (e.g., is change detection the primary or secondary task? Are the stimuli simple or naturalistic and complex? Does the experiment happen in the laboratory or in real-life situations?).

The first proposal, *overwriting*, states that information from the first image is indeed encoded (i.e., a representation of the scene is abstracted), but will afterwards be either replaced or overwritten by information from the second scene. When presented with complex scenes (e.g., naturalistic stimuli, such as photographs or video), some visual information is encoded into a representation of the scene; however, details of the scene that were not attended to are replaced by information from the subsequent scene (after change is introduced). This model predicts, therefore, that change blindness would happen exclusively for unattended details of the visual scene.

The *first impressions* model claims that the main purpose of perception is to extract meaning from our environment (and, consequently, visual scenes). As such, features of a scene that more adequately inform us of its meaning are encoded; everything else, being peripheral to this goal, goes unchecked and remains absent from our representation of the scene. Contrary to the *overwriting* model, details of the first scene are in fact encoded – but only those that communicate meaning. It is the details of the subsequent scene, including the change, that are not processed. Note that this model, to work as stated, requires that the meaning of the scene is not changed by the introduction of a change.

A third model proposes that *everything is stored*, *though not always compared*: central to this model is the assumption that we are able to simultaneously hold two representations of a visual scene that contain differences (changes) between them. Change blindness occurs as these differences, until pointed out, remain undetected. This "pointing out" can be, for instance, a change that either draws attention due to its perceptive properties, or that alters the meaning of the scene significantly. If such highlighting of the change does not happen, "participants may form representations of the details both before and after the change, but never bother to compare the two" (Levin, Simons, Angelone, & Chabris, 2002, p. 300). Studies aiming to provide evidence of the existence of both pre- and post-change representations of visual scenes have simultaneously provided evidence against an interpretation of change blindness as occurring due to *nothing being stored* – in short, the world itself, being ever-present, is our storage of scenes we witness. Such studies adopt the incidental change task; when participants fail to detect a changing object (e.g., disappearing), they are asked if they remember that object. This was the method used by Simons, Chabris, Schnur, and Levin (2002): in a real-life interaction, participants failed to notice when a basketball would disappear from the interlocutor's hands, during a disruption by a passing group of pedestrians. However, they would later recall the basketball, even describing it correctly. In a laboratory setting Mitroff, Simons, and Levin (2004) demonstrated that participants could recognize an object present in a *prechange* conditions (i.e., replaced by a different object during the change), even when they had no conscious of that change.

Experimental paradigms

Learning about the circumstances in which we fail to detect changes also informs us as to how and when we do detect them; as such, part of the literature uses the term *change detection*, referring to "the visual processes involved in first noticing a change" (Rensink, 2002, p. 246). Whether under the name *change blindness* or *change detection*, many paradigms have been proposed to shed light on the phenomenon. The earlier studies on change detection were not yet using the naturalistic stimuli (photography or video) that are commonplace today: French (1953), for instance, used patterns of dots presented to participants with a 6-second interval. Participants were asked to determine if the two patterns were the same (or if something had changed). This method falls under the category of *gap-contingent* methods: change happens during a period between two stimuli, in which nothing is shown (in French's case, the 6 seconds between patterns). The time between stimuli is today referred to as an *interstimulus interval* (hereinafter referred to as ISI). Other studies picked up on this method (e.g. Pashler, 1988); a particular variation went beyond showing the two stimuli once and asking for a response (a method known as "forced response"), and instead it would rapidly cycle the two stimuli and a blank screen (ISI) in alternation until observers detected a change – the *flicker paradigm* (Rensink, O'Regan, & Clark, 1997; Simons, 2000). In this paradigm, the change happens repeatedly until either the observer detects it, or a set number of cycles is reached and the trial ends (see Figure A2). Unlike forced response paradigms, the measured variable is not the number of correct or incorrect responses, but reaction times – or how long before the participants noticed the change.

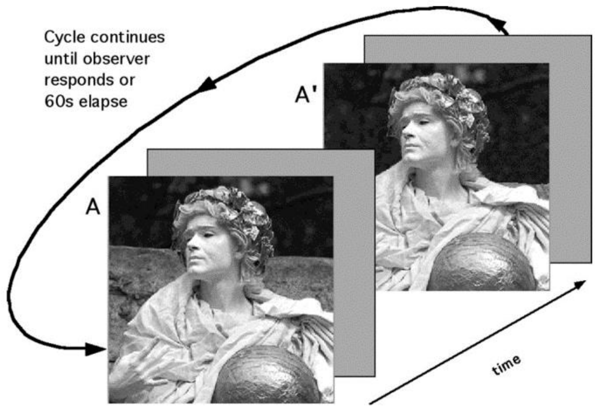


Figure A2. The flicker paradigm (Rensink et al, 1997). Original image (A) and modified (A´, with lowered wall) are alternated with a blank screen. Cycle repeats a fixed number of times, or until participants stops it.

Note. Reprinted from "Changes", by Rensink. R., 2002, *Progress in brain research, 140*, p. 199. Copyright 2002 by Elsevier Ltd.

A blank screen is not the only way to disrupt the viewing of a scene and introduce a change. Other noteworthy methods include using "*mudsplashes*" – instead of replacing the entire image, shapes are momentarily superimposed while changes are introduced (O'Regan, Rensink, & Clark, 1999) – or using our eye's natural movements, whether introducing changes during a saccade (e.g., Henderson & Hollingworth, 1999), or during natural and induced blinking (e.g., O'Regan, Deubel, Clark, & Rensink, 2000). In all these methods, severe impairment in change detection is observed.

The increase in usage of more realistic materials continued until the experiment was taken out of the lab, with designs that simulate real life scenarios. Levin and Simons (1997) composed short videos in which changes were introduced during *cuts* (a jump to a different camera angle). Changes were not all minimal (e.g. a random object in the background), sometimes changing the actors themselves (the attended object); nevertheless, they went largely unnoticed by participants. The same authors (Simons & Levin, 1998) finally took to the streets, replacing an actor during two real-life interactions: pedestrians were asked for directions by one of the experimenters, but were interrupted by two men carrying a door between them. While behind the door, the experimenter is replaced by a second actor. In two studies, the change was noticed by seven out of 15 participants, and four out of 12 participants.

Conclusions from these studies seem to converge in two ways:

a) our capacity to retain detailed information regarding a visual scene is, at most, limited – a disruption in a visual field (such as eye movements, blinks, or blank screens) seems to impair our ability to retain information from one moment to the next (Simons, 2000);
b) the role of one key factor in change blindness performance – attention. When Rensink and collaborators (1997, p. 372) established the flicker paradigm, they also concluded that detection of change "occurs only when that object is given focused attention", supporting it with results showing that change detection was better in areas of "central interest" (defined by the authors as those objects or regions most frequently named in descriptions of the observed scene).

On attention

Attention, in its everyday meaning, is a concept familiar to everyone. William James (1890, p. 403) had no issue in stating it directly – "Every one knows what attention is" – but then proceeds to define it as "the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought." A full exploration of attention and its impact in cognition is beyond the scope of this literature review, but one particular aspect is relevant to change blindness/change detection: its limits, and how they are efficiently managed to respond to the situation at hand. Harris and Jenkin (2001), in their thorough definition of attention and its different types, begin by highlighting how it is a compromise: it is a finite resource to be allocated to some stimuli, in detriment of others. It is a second filter after what the senses already impose (e.g., the limits of our field of vision). In short, no individual can attend to everything.

The role of attention on change detection was expanded on by other studies. The already mentioned studies of change blindness using film (Levin & Simons, 1997, p. 505) showed that "even dramatic changes to the very object that is the center of attention often go unnoticed" – attention might be undeniably important, but it is not sufficient to detect changes. Simons and Chabris (1999) focused on this "inattentional blindness" (see Mack & Rock, 1998, for more on this concept) – in their method, a video task in which two teams (dressed either in black or white shirts) passed a basketball between them. Participants were instructed to count how many passes were done by one team. During the video, a man dressed as a gorilla would walk across the image – in one condition, the gorilla was highly salient (even facing the camera and thumping its chest) and remained in sight for nine full seconds. Even in such conditions, approximately half of the sample failed to detect the change. Note that the object (the man dressed as a gorilla) would often cross the region that was being attended, and still go unnoticed. It is important to highlight that Simons and Chabris' (1999) task did not instruct participants to detect change at the start of the experiment, instead giving them a different task to perform: counting the passes of a specific team. In other words, diverting attention to a task (and away from the introduced change) impaired change detection despite it happening in the attended region of the scene. This is in accordance with Rensink's (2002) suggestion that instead of creating a detailed representation of visual scenes, we rely on a more adapted representation, adequate to the specific demands of each task, and the participant's expectations – a *dynamic representation*.

In short, when we are involved and focused on a demanding task, an object change "does not automatically capture attention when it is incidental to the primary focus of attention" (Simons & Mitroff, 2001, p. 189). This proposal is also compatible with something Simons and Chabris (1999) noted from their experiment: that the change was *unexpected*.

On expectations

Change detection of an unexpected object is the core feature of the "incidental approach" to the topic in research – as opposed to an "intentional approach" (Simons & Mitroff, 2001). Studies such as French (1953), and Rensink and colleagues (1997), instruct participants to try and detect a change in a visual scene – that is their main task – and as such are examples of the intentional approach to change detection. Stimuli are not always naturalistic, and the experiments are usually conducted in a laboratory setting. Moreover, while intentional visual search tasks are also present in our day-to-day life (e.g., trying to locate an empty spot in a car park, or a coin that has been dropped on the sidewalk), they do not represent every mode or method of visual processing while we navigate rich visual environments.

The incidental approach assumes change detection as secondary, instructing observers to perform a different task while the change is implemented in the scene. This is the category in which studies such as Levin and Simons (changes between cuts in motion pictures; 1997), and Simons and Chabris (gorilla in the background; 1999) would fall under: what is measured is our performance in detecting changes that are secondary to the task at hand. Using naturalistic stimuli and environments, the incidental approach aims to generalize findings to real-life situations in which we are not intentionally looking for a change.

Expectations, however, go beyond whether we expect a change to happen. It is hard to even define if we expect changes when we go about our days: we certainly don't think that everything in our visual field is immutable, and we expect such mutability (e.g., we know outdoor ads can flash different colors, even if we cannot always predict them), but we have also learned not to expect *every* change (e.g., an outdoor ad becoming sentient and walking away). Our surrounding, itself, is rich in cues regarding what we can expect. For instance, someone yelling out the lyrics to a song might be expected in a rock concert but would certainly hold our attention at the opera; a pedestrian might attract little attention if she is simply crossing the road but might be more noticeable if doing it while floating six meters above the ground.

How does this translate to research? What are, then, the types of expectations that we abstract from context?[1] To be able to answer these questions it is important to establish that we do abstract something from context – and change blindness tasks have also proven helpful here. Recall Rensink and colleagues (1997, p. 368), and their study in which the flicker paradigm was introduced; one of the conclusions was that object identification was faster when objects were deemed by participants as "important to the scene". It is not surprising that subjects can determine what is of central or marginal interest in a scene: a meaningful representation of a visual scene can be obtained in a glance as short as 150 milliseconds (Biederman, 1981). This proposition that we can quickly obtain a *gist* of a visual scene found support from different studies: for instance, changes that alter the gist of a scene are more rapidly detected (Sampanes, Tseng, & Bridgeman, 2008); scene inversion (i.e., showing pictures upside down) eliminates the effect of object interest, and no benefit for object of central interest is registered in the form of significantly lower reaction times (Kelley, Chun, & Chua, 2003; Shore & Klein, 2000); scene jumbling (method by which an image is cut into various pieces that are randomly recombined) and randomly rearranging objects also translates to impairments in change blindness performance, suggesting that an understandable

[1] Context informs many other cognitive processes that fall outside the scope of this literature review and project (for a more broad review, see Chun, 2000). As such, this section focuses on contextual influences that relate directly to change blindness.

scene context can help in change detection (Biederman, Glass, & Stacy, 1973; Zimmermann, Schnier, & Lappe, 2010).

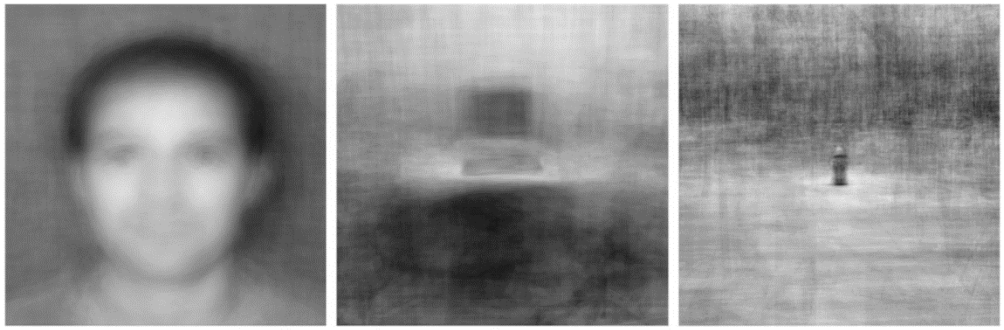


Figure A3. Average of hundreds of pictures containing the same object, centered and scaled to match. The most common features of the object's surrounding space are represented with higher intensity, showing how objects establish frequent spatial relations.

Note. Reprinted from "The role of context in object recognition", by Oliva, A. and Torralva, A., 2007, *Trends in cognitive science, 11(12)*, p. 521. Copyright 2002 by Elsevier Ltd.

If we perceive the gist (i.e., general meaning, or the essence) of a visual scene, we may also be able to draw helpful expectations regarding, for instance, what objects might be present in the image, what absolute positions they may occupy (e.g., top or bottom), or even their spatial relationship to other objects in the scene. When observing a scene, we define, then, "useful constraints on the range of possible objects that can be expected to occur within that context" – the act of abstracting such constraints from context defined the *contextual cueing paradigm* (Chun, 2000, p. 171). The helpfulness of the information we draw from context derives from the fact that certain types of object usually appear together in our visual field – the correlations, or how objects cluster, are not random, instead co-varying according to specific environments and generating predictability (and we seem to learn these co-variations both fast and automatically; Fiser & Aslin, 2001, 2005). For instance, in most situations, a desk is usually accompanied by a chair; Oliva and Torralba (2007, p. 521) illustrate the consistency of these spatial relationships and configurations by overlapping hundreds of images that represent, in its natural context, a specific object (edited to have the same scale and pose in each picture). The result shows patterns that indicate how objects are usually organized around the object, revealing its "influence beyond its own boundaries" (see Figure A3). The same can be said for an object's position: in a photograph of a landscape, we expect a house to be on the bottom half.

Research has backed up the assumptions of the contextual cueing paradigm. For example, Biederman, Mezzanotte, and Rabinowitz (1982) show that when objects in a scene violate the typical spatial relationships (e.g., a floating sofa, a fire hydrant that appears on top of a mailbox), its *detectability* is reduced – in other words, higher reaction times are registered for when the object is correctly detected. One particularly interesting violation was named *Probability* by the authors and defined as "the likelihood of a given object being in a given scene"; they add that being from a semantic nature, it would require "access to the referential meaning of the object" (p. 146). Semantic processing of cues from context was also explored by Palmer (1975), in an experiment in which specific contexts were initially shown to participants (e.g., depictions of a kitchen or a living room), followed by illustrations of object that were presented between 20 and 120 milliseconds. Results show that independently of exposure time, "the probability of correctly identifying an object is highest following an appropriate context" (p. 521). More recent research efforts, while resorting to different methods, report similar conclusions: objects are correctly identified more frequently if placed in a *consistent* background (i.e., one that depicts where they naturally occur in the real-world), and a short glimpse at the visual scene (80 milliseconds) is enough for the effect to be observed (Davenport & Potter, 2004; Munneke, Brentari, & Peelen, 2013)[2].

The research reviewed above offers empirical support for the key role of context in perception, focusing mainly on identification and using a variety of methods. Change blindness and change detection can, however, happen without our being able to identify the change or changing object (Sagi & Julesz, 1984; Straube & Fahle, 2011; Xu, 2008). We can perceive changes to a visual scene despite not being able to pinpoint exactly what changed – *detection tasks* instruct the participants to do just that. At this stage, the question to answer is: does consistency (or congruency) between changed object and scene context, in change blindness tasks, benefit *detection*? Hollingworth and Henderson (2000), for instance, note the opposite: a main effect of semantic consistency in which responses to inconsistent targets register significantly faster reaction

[2] Note: these studies were referenced to illustrate contextual cueing – due to differences in method, these studies cannot be classified as change blindness studies, or in the line of the flicker paradigm.

times, when compared to consistent targets. In short, “a change to an object is more easily detected when that object is semantically inconsistent with its scene” (p. 234).

Hollingworth and Henderson (2000) used the flicker paradigm proposed by Rensink et al (1997), which is named by LaPointe and Milliken (2016, p. 8) as the “one task that has produced performance benefit for semantically incongruent object reliably and robustly”. Not surprisingly, they used the same task to show that eye-movement is drawn to such objects.

LaPointe, Lupianez, and Milliken (2013) attempted to clarify the impact of task type: detection vs. identification. To do so, the authors used a close variation of the flicker task (in fact, a revision of the authors’ version was used in this dissertation) with original stimuli. In the first experiment, the same benefit for incongruent trials was observed (in the form of lower reaction times), but it is debatable if the task can be categorized as a *detection* task when after each trial participants were “prompted to make a one-word response to identify the changing object” (p. 105). There is no doubt that, if participants were unaware of this specific part of the task at the start, the first trial might be a true detection task. But participants might also assume, after the first trial, that they would have to *identify* objects from then onwards, triggering processes typical of identification tasks.

Experiments 2 and 3 yield results that might be more helpful in highlighting a possible difference between task type, as both studies include a condition in which the ISI (interstimulus interval) is eliminated (i.e., 0ms, meaning that the original and edited image would cycle between them with no blank screen in between). Experiment 2 maintained the requirement of identifying the changing object, and faster reaction times were registered for congruent trials. However, experiment 3 asked participants not to identify the object, but to indicate the location of the change (which does not imply identification) – for both the 250 milliseconds and 0 milliseconds ISI conditions, a benefit to incongruent trials was observed. In short, an identification requirement seems to benefit congruency, while a detection requirement prioritizes incongruency. This conclusion, consistent with the literature reviewed so far, is also supported in three experiments by Ortiz-Tudela et al (2016), in which both detection and identification are analyzed: detection favors the incongruent, while identification benefits the congruent.

Why does it happen? LaPointe and Milliken (2016, 2017) suggest an explanation based on a dual-process model. On one side, and following the suggestions of contextual cueing, they emphasize the role of factors such as prior knowledge or expertise in identification tasks (by means of abstracting expectations regarding objects from context information). As for detection, it benefits incongruency due to “attention capture processes”, fine-tuned to details considered informative, conflicting, and therefore incongruent in their contexts.

On stereotypes

The leading question of this project picks up on the idea that context in a visual scene (i.e., a stimulus background in a flicker task) can trigger expectations regarding what objects are probably present or absent. We ask if stereotypes can be seen as context, and if so, what effect would be observed if they were included in a change detection task? The question is based on the idea that, much like the context mentioned in the reviewed literature, stereotypes are known to activate other concepts.

Asserting a definition of stereotype is useful. Although the term tends to assume a negative connotation in everyday language (as they are equated with their most negative and visible consequences), studies in social cognition opt to underline that they are in fact a product of human nature – as a phenomenon, it is heuristic and efficient. Greenwald and Banaji (1995), for instance, define it as a group of socially shared beliefs, regarding traits that are characteristic of members of a particular social category. In the same review, they quote Allport (1954, p.191): “A stereotype is an exaggerated belief associated with a category”.

The difference between common use of the term *stereotype* and its academic definition corresponds, then, to the difference between an inadequate tool, and the unfortunate use of a valuable tool (for a historical review of stereotypes under the lens of social cognition, see Garcia-Marques & Garcia-Marques, 2003, and Monteith, Woodcock, & Gulker, 2013). Note that tools, throughout humankind’s history, have been invented in the interaction with our environments, so that we could more easily navigate it, manipulate it, ever more efficiently. Macrae and colaborators (1994) underline this parallel between physical tools and our cognitive toolbox: a collection of various mechanisms that allow for an efficient processing of the worlds excess of stimuli. In their words, “individuation, in its many guises, is a rather time consuming and effortful affair” (p. 37). In seminal work on stereotypes, the authors underlined their energy-saving potential. Participants were asked to perform two simultaneous tasks: forming impressions from a list of trait descriptors while attending to information in audio form. Trait lists included either a name (e.g., John), or a name accompanied by a stereotypical label (e.g., John – *skinhead*). As hypothesized, the presence of stereotypical information simplified the impression forming process, consequently freeing resources to the second task, and improving performance. This demonstrates that the effects of stereotypes go beyond the usually negative outcomes, such as discrimination.

In the popular definition, stereotypes have a negative connotation partly (if not entirely) due to the visibility of some behaviors that stem from them, and the negative consequences of such behaviors. As a result, stereotypes are often equated with concepts such as prejudice or discrimination, although they are, by definition, different. Prejudice refers to an attitude towards a member of a group by the mere fact that he or she belongs to that group; it is generally a negative attitude – in other words, a negative evaluation of a person based on perceived belonging to a social category or group. Discrimination, on the other hand, refers to behavior, or taking (again, generally negative) action towards someone based only on perceived belonging to a group.

Other misconceptions regarding stereotypes can be named. For instance, while the most visible and talked about stereotypic beliefs are false (e.g., Latinos are lazy), that is not necessarily true in every situation. The stereotype that men are more aggressive than women can be demonstrated as truthful by crime rates and records of domestic violence, in which the perpetrators are mostly male (Portugal included, PORDATA, 2018).

It is also a misconception that stereotypes always translate into behaviors. For instance, Devine's (1989) dualist model predicts that while stereotype activation is indeed automatic, it is possible to override its influence. Given enough time, controlled processes can allow for non-stereotypical personal beliefs to inhibit the effect of stereotypes. In the author's results, this was true for participants who scored low on the modern racism scale (McConahay, Hardee, & Batts, 1981). In short, even when we are aware of the contents of a stereotype, and even when that stereotype is automatically activated, its impact can be overridden by personal beliefs that contradict it.

Further research has shown that even this automaticity of stereotype activation cannot be taken for granted (see Devine & Sharp, 2009; Monteith, Woodcock, & Gulker, 2013), instead being dependent on a series of conditions. For instance, "individuals who are preoccupied with other matters tend to not experience automatic stereotypes" with unattended social cues failing to activate social categories (Devine & Sharp, 2009, p. 65). Gilbert and Hixon's (1991) influential study suggested that, when participants were kept cognitively busy by a task (e.g., memorizing a number), stereotype activation failed. In a different experiment, activation was permitted: subjects who were afterward kept busy with a visual search task were more likely to use those stereotypes while forming impressions. Cognitive busyness, then, might disable the automatic activation of stereotypes, but when activation occurs, busyness amplifies their usage (see also Collier & Shaffer, 1999).

But even when undistracted, group membership and stereotypes are only conditionally activated. For example, it is necessary that the perceived person is encoded in a socially meaningful way – using the words of Macrae, Bodenhausen, Milne, Thorn, and Castelli (1997), it depends on *processing goals*. Summarizing their work: participants saw faces of women with the instruction to either detect the presence of a white dot (feature-detection), or to determine if an animate or inanimate object was shown (semantic-judgement). In a subsequent lexical decision task containing both female-stereotypic and counterstereotypic words, participants in the semantic-judgement condition were faster at classifying stereotypic terms. This effect of semantic processing (perceiving people as socially relevant stimuli) is even sensitive to eye gaze (Macrae, Hood, Milne, Rowe, & Mason, 2002).

The same stimulus, such as the same photo of a person, might not even activate stereotypes consistently, depending on the context within which it is presented. As shown by Wittenbrink, Judd, and Park (2001), different background pictures (e.g., stereotypically positive: church interior vs. stereotypically negative: street corner) interfered with participants' "automatic responses to social category cues" (p. 823).

While a full account of all factors interfering with stereotype activation would fall outside the relevant range of this review, an example of a situational influence can be useful in demonstrating the variety of ways in which stereotype activation is, indeed, conditional. In a study by Sinclair and Kunda (1999, p. 855), inhibition of "applicable stereotypes" is shown to be motivated by self-protection: participants are more likely to inhibit stereotypes associated with black individuals when praised by one (when compared to those criticized by a black judge). This study, along with the effects of context (Wittenbrink et al., 2001), demonstrate that the variability in which conditions might determine the activation or inhibition of stereotypes permeate our everyday lives and the complex social world.

There is reason to expect a stereotypical stimulus (e.g., a person, member of a group) to activate several other concepts. Theory on implicit attitudes and their manifestations (see

Greenwald, Rudman, Farnham, Nosek, & Mellott, 2002) tells us that social knowledge is structured in an associative network of concepts that associate between themselves with different *strength*. Some definitions are relevant:

- *Concept*: representations of people, groups, attributes – these being particularly relevant as they can assume positive or negative valence;
- *Associations*: relationships between concepts;
- *Strength of association*: can be understood as the potential of a concept to active other concepts that it is associated with or connected to; associations are bilateral.

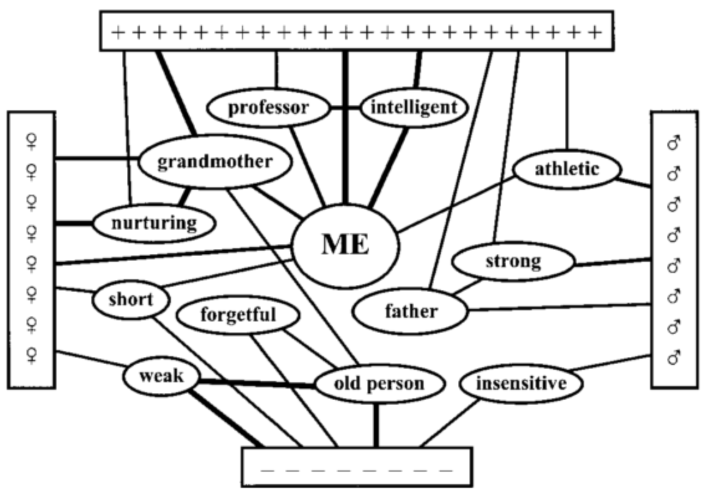


Figure A4. Social knowledge structured as an associative network, in which each concept (elipses), once activated, can activate neighboring concepts. Line thickness represents strength.

Note. Reprinted from "A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept", by Greenwald et al, 2002, *Psychological Review, 109(1),* p. 5. Copyright 2002 by American Psychological Association, Inc.

Figure A4 exemplifies this by depicting concepts that surround the self, but are also more or less associated with gender, as they are placed to the right (male), or the left (female). The height of each concept indicates valence. From the same example, we can say that by activating the concept of *weak*, the concept of *old person* might also be activated (e.g., more accessible in our memory) – this is, in fact, one of the most frequent paradigms in studies of implicit social cognition: *priming*, listed as one of the "major routes to stereotype maintenance" by Hilton and Hippel (1996, p. 248) on their text on stereotypes. In that text, the authors review the multiple ways in which priming has been a staple of research in social cognition and stereotypes, asserting that "priming plays a dramatic role in the perception of out-group members" (p. 248).

We know, then, that our implicit definitions of social categories and their members create expectations – usually in what qualities or behaviors we expect from those individuals. In the realm of change blindness, this would translate to activating and expecting certain objects in a context that contains a stereotype. For instance, if we perceive a man in a bedroom, might we not expect to see a pair of men's shoes? If we see a woman in that bedroom, would a dress not be a more congruent object?

Other factors that influence change blindness

While it would be outside the scope of this literature review to identify all the factors known to impact change blindness in some way, naming some of the most relevant will contribute to a better overall understanding of the phenomenon.

Age, for example, is one such factor. For instance, Rizzo et al. (2009, p. 252) note how change blindness increases with both age and the onset of Alzheimer's disease – an increase that correlates negatively with "performance on cognitive tests that depend on attention, memory, and executive functions". Caird, Edwards, Creaser, and Horrey's (2005) study also stresses the effects of aging by using a modified flicker task in which the correct decision implies the detection of an appearing-disappearing object. The stimuli simulate real-life scenarios encountered while driving – namely, intersections. Younger participants performed better than older participants, with significantly more correct decisions.

With driving, it is common knowledge that experience is required to develop the necessary skills. In research, this translates into the influence of expertise and familiarity. Werner and Thies (2000), for example, show the benefit of domain-specific expertise by comparing the performance of football players and non-players in a change detection task involving football-related and unrelated stimuli. Results clearly show a performance benefit for experts, when compared to novices.

While expertise in football playing might come from conscious effort, training, and dedication, not all expertise comes about so declaratively. As Jones et al. (2003) show, heavy social users of cannabis and alcohol are faster at detecting substance-related changes (in a flicker task) when compared to substance-neutral changes. In a different example, Buttle and Raymond (2003) report significantly improved performance in detecting changing faces when faces are famous and previously known (they call this the *superfamiliarity effect*).

The example of faces has also been used to demonstrate a cultural influence in change blindness. In Humphreys, Hodsoll, and Campbell's (2005) adaptation of the flicker task, White Caucasian and Indian Asian participants detected changes faster when they occurred in the faces of people of the same ethnicity as theirs. The authors discuss their results as a possible outcome of other known processes; for example, differential encoding of faces due to different amounts of exposure, or a tendency to see the outgroup as more homogenous than the ingroup.

These studies often rely on the flicker task, or an adaptation of it. But we have also reviewed studies that used video stimuli, as well as real-life scenarios. Gibbs, Davies, and Chou (2016, p. 17), in a systematic review of studies with such methods, found three familiar factors affecting change blindness performance: "increasing attention, the saliency of the changed object and spatial violations significantly reduce CB".

Saliency, being ubiquitous in every task involving visual stimuli, should be addressed more thoroughly. It refers to the prominence of any stimulus, or a particular feature of a stimulus, that stands out from the rest (Gibbs et al., 2016). In practical terms, it means it captures human attention – in visual search tasks, this equates to a higher probability of holding our gaze, or fixation. This is not to say that saliency is dependent on the stimulus alone, capturing our attention passively without the observing subject's intervention. Instead, it works similarly to how we perceive color: it is not a quality of the object itself, depending also on the qualities of incident light, and the peripheral sensory equipment of the observer. As such, an object is salient when it has certain properties (e.g., color, size, contrast), but also when it is relevant to the observer, and in accordance with situational characteristics. A bright red smartphone ad might go completely unnoticed by a human observer, but it might also mobilize our attention due to color and size. The same ad might be more salient to an observer looking to acquire a new smartphone. However, even that interested observer might miss the ad if it blends with an equally bright, equally colorful casino wall. In short, a certain "biological relevance" is needed (Jenkin & Harris, 2001). Salience happens when both observer, stimulus features, and situational demands meet halfway and interact.

In research, those stimulus properties are referred to as *bottom-up*, or low-level saliency, interacting with early-stage visual processing. On the other hand, high-level saliency (*top-down*) includes semantic or motivational processing of objects and scenes (e.g., contextual cueing, congruency effects). The coexistence of these processes makes it so that the impact of saliency in change blindness performance is not easy to predict – and the literature on the subjects reflects just that. For instance, Boyer et al. (2011) used the flicker paradigm to implement changes either in high-saliency or low-saliency objects (as determined by an algorithm, Itti, Koch, & Niebur, 1998), and reported faster detection for high-saliency changes. However, this influence of bottom-up processes seems to depend on the nature of the task: in free-viewing, preparing for a memory task, salient and more complex objects prevail at attracting fixation; but when instructed to search for a specific low-saliency object, nontarget saliency has negligible impact (Underwood & Foulsham, 2006; Underwood, Foulsham, van Loon, Humphreys, & Bloyce, 2006). In fact, one factor that tends to override low-level saliency is congruency between objects and background, with incongruent changes being detected faster independently of how salient the features of the changing object are.

References

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (1st ed., pp. 213–233). London: Routledge.

Biederman, I., Glass, A. L., & Stacy, E. W. (1973). Searching for objects in real-world scenes. *Journal of Experimental Psychology*, *97*(1), 22–27. https://doi.org/10.1037/h0033776

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detection and judging objects undergiong relational violations. *Cognitive Psychology*, *14*(2), 143–177. https://doi.org/10.1016/0010-0285(82)90007-X

Boyer, T. W., Smith, T. G., Yu, C., & Bertenthal, B. I. (2011). Visual attention and change detection. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1735–1740).

Buttle, H., & Raymond, J. E. (2003). High familiarity enhances visual change detection for face stimuli. *Perception & Psychophysics*, *65*(8), 1296–1306. https://doi.org/10.3758/BF03194853

Caird, J. K., Edwards, C. J., Creaser, J. I., & Horrey, W. J. (2005). Older driver failures of attention at intersections: Using change blindness methods to assess turn decision accuracy. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *47*(2), 235–249. https://doi.org/10.1518/0018720054679542

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, *4*(5), 170–178. https://doi.org/10.1016/S1364-6613(00)01476-5

Collier, C. A., & Shaffer, D. R. (1999). Activation and use of racial stereotypes in personnel decisions: A test of two theories. *Journal of Applied Social Psychology*, *29*(11), 2292–2307.

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*(8), 559–564. https://doi.org/10.1111/j.0956-7976.2004.00719.x

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5–18. https://doi.org/10.1037/0022-3514.56.1.5

Devine, P. G., & Sharp, L. B. (2009). Automaticity and control in stereotyping and prejudice. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and descrimination* (pp. 61–135). Hove, East Sussex: Psychology Press.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*(6), 499–504. https://doi.org/10.1111/1467-9280.00392

Fiser, J., & Aslin, R. N. (2005). Encoding multielement scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General*, *134*(4), 521–537. https://doi.org/10.1037/0096-3445.134.4.521

French, R. S. (1953). The discrimination of dot patterns as a function of number and average separation of dots. *Journal of Experimental Psychology*, *46*(1), 1–9.

Garcia-Marques, L., & Garcia-Marques, T. (2003). Mal pensa quem não repensa: Introdução ao estudo dos estereótipos sociais numa perspectiva cognitiva. In T. Garcia-Marques & L. Garcia-Marques (Eds.), *Estereótipos e a sua influência no processamento de informação* (1st ed., pp. 11–25). Lisboa: Instituto Superior de Psicologia Aplicada.

Gibbs, R., Davies, G., & Chou, S. (2016). A systematic review on factors affecting the likelihood of change blindness. *Crime Psychology Review*, *2*(1), 1–21. https://doi.org/10.1080/23744006.2016.1228799

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, *60*(4), 509–517. https://doi.org/10.1037/0022-3514.60.4.509

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4–27. https://doi.org/10.1037/0033-295X.102.1.4

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitdes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*(1), 3–25. https://doi.org/10.1037//0033-295X.109.1.3

Harris, L. R., & Jenkin, M. (2001). Vision and attention. In M. Jenkin & L. R. Harris (Eds.), *Vision and attention* (pp. 1–17). New York: Springer.

Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science*, *10*(5), 438–443. https://doi.org/10.1111/1467-9280.00183

Hilton, J. L., & Hippel, W. Von. (1996). Stereotypes. *Annual Review of Psychology*, *47*, 237–271.

Hollingworth, A., & Henderson, J. M. (2000). Semantic informativeness mediates the detection of changes in natural scenes. *Visual Cognition*, *7*(1–3), 213–235.

https://doi.org/10.1080/135062800394775

Humphreys, G. W., Hodsoll, J., & Campbell, C. (2005). Attending but not seeing: The “other race” effect in face and person perception studied through change blindness. *Visual Cognition*, *12*(1), 249–262. https://doi.org/10.1080/13506280444000148

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259.

James, W. (1890). The principles of psychology. *New York Holt*, *1*, 697. https://doi.org/10.1037/10538-000

Jenkin, M., & Harris, L. (2001). Vision and attention. In M. Jenkin & L. Harris (Eds.), *Vision and attention* (pp. 1–18). Springer.

Jones, B. T., Jones, B. C., Smith, H., & Copley, N. (2003). A flicker paradigm for inducing change blindness reveals alcohol and cannabis information processing biases in social users, 235–244.

Kelley, T. A., Chun, M. M., & Chua, K.-P. (2003). Effects of scene inversion on change detection of targets matched for visual salience. *Journal of Vision*, *3*(1), 1–5. https://doi.org/10.1167/3.1.1

LaPointe, M. R. P., Lupianez, J., & Milliken, B. (2013). Context congruency effects in change detection: Opposing effects on detection and identification. *Visual Cognition*, *21*(1), 99–122. https://doi.org/10.1080/13506285.2013.787133

LaPointe, M. R. P., & Milliken, B. (2016). Semantically incongruent objects attract eye gaze when viewing scenes for change. *Visual Cognition*, *24*(1), 1–15. https://doi.org/10.1080/13506285.2016.1185070

LaPointe, M. R. P., & Milliken, B. (2017). Conflicting effects of context in change detection and visual search: A dual process account. *Canadian Journal of Experimental Psychology*, *71*(1), 40–51. https://doi.org/10.1037/cep0000105

Levin, D. T., Drivdahl, S. B., Momen, N., & Beck, M. R. (2002). False predictions about the detectability of visual changes: The role of beliefs about attention, memory, and the continuity of attended objects in causing change blindness blindness. *Consciousness and Cognition*, *11*(4), 507–527. https://doi.org/10.1016/S1053-8100(02)00020-X

Levin, D. T., Momen, N., Drivdahl, S. B., & Simons, D. J. (2000). Change blindness blindness:

The metacognitive error of overestimating change-detection ability. *Visual Cognition*, *7*(1–3), 397–412. https://doi.org/10.1080/135062800394865

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin and Review*, *4*(4), 501–506. https://doi.org/10.3758/BF03214339

Levin, D. T., Simons, D. J., Angelone, B. L., & Chabris, C. F. (2002). Memory for centrally attended changing objects in an incidental real-world change detection paradigm. *British Journal of Psychology*, *93*(3), 289–302. https://doi.org/10.1348/000712602760146224

Mack, A., & Rock, I. (1998). An overview. In A. Mack & I. Rock (Eds.), *Inattentional blindness* (pp. 21–26). London: The MIT Press.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., Thorn, T. M. J., & Castelli, L. (1997). On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Social Psychology*, *33*(5), 471–489. https://doi.org/10.1006/jesp.1997.1328

Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (2002). Are you looking at me? Eye gaze and person perception. *Psychological Science*, *13*(5), 460–464. https://doi.org/10.1111/1467-9280.00481

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, *66*(1), 37–47. https://doi.org/10.1037/0022-3514.66.1.37

McCallum, R., & Lucas, G. (1999). *Star Wars. Episode I, The Phantom Menace*. United States: 20th Century Fox.

McConahay, J. B., Hardee, B. B., & Batts, V. (1981). Has racism declined in america? It depends on who is asking and what is asked. *Journal of Conflict Resolution*, *25*(4), 563–579. https://doi.org/10.1177/002200278102500401

Mitroff, S. R., Simons, D. J., & Levin, D. T. (2004). Nothing compares 2 views: Change blindness can occur despite preserved access to the changed information. *Perception and Psychophysics*, *66*(8), 1268–1281. https://doi.org/10.3758/BF03194997

Monteith, M. J., Woodcock, A., & Gulker, J. E. (2013). Automaticity and Control in Stereotyping and Prejudice: In D. E. Carson (Ed.), *The Oxford handbook of social cognition* (pp. 74–94). New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199730018.013.0005

Munneke, J., Brentari, V., & Peelen, M. V. (2013). The influence of scene context on object recognition is independent ofattentional focus. *Frontiers in Psychology*, *4*(August), 1–10. https://doi.org/10.3389/fpsyg.2013.00552

O'Regan, J. K., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition*, *7*(1–3), 191–211. https://doi.org/10.1080/135062800394766

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of "mudsplashes." *Nature*, *398*(4), 34. https://doi.org/10.1038/17953

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, *11*(12), 520–527. https://doi.org/10.1016/j.tics.2007.09.009

Ortiz-Tudela, J., Milliken, B., Botta, F., LaPointe, M., & Lupiañez, J. (2017). A cow on the prairie vs. a cow on the street: Long-term consequences of semantic conflict on episodic encoding. *Psychological Research*, *81*(6), 1264–1275. https://doi.org/10.1007/s00426-016-0805-y

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*(5), 519–526. https://doi.org/10.3758/BF03197524

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*(4), 369–378. https://doi.org/10.3758/BF03210419

PORDATA. (2018). PORDATA - Estatísticas, gráficos e indicadores de Municípios, Portugal e Europa. Retrieved from http://www.pordata.pt/

Rensink, R. A. (2002a). Change detection. *Annual Review of Psychology*, *53*, 245–277. https://doi.org/10.1146/annurev.psych.53.100901.135125

Rensink, R. A. (2002b). Changes. *Progress in Brain Research*, *140*, 197–207. https://doi.org/10.1016/S0079-6123(02)40051-9

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368–373. https://doi.org/10.1111/j.1467-9280.1997.tb00427.x

Rizzo, M., Sparks, J., McEvoy, S., Viamonte, S., Kellison, I., & Vecera, S. P. (2009). Change blindness, aging, and cognition. *Journal of Clinical and Experimental Neuropsychology*, *31*(2), 245–256. https://doi.org/10.1080/13803390802279668

Sagi, D., & Julesz, B. (1984). Detection versus discrimination of visual orientation. *Perception*, *14*(5), 619–628. https://doi.org/10.1068/p130619

Sampanes, A. C., Tseng, P., & Bridgeman, B. (2008). The role of gist in scene recognition. *Vision Research*, *48*(21), 2275–2283. https://doi.org/10.1016/j.visres.2008.07.011

Shore, D. I., & Klein, R. M. (2000). The effects of scene inversion on change blindness. *Journal of General Psychology*, *127*(1), 27–43. https://doi.org/10.1080/00221300009598569

Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition*, *7*(1–3), 1–15. https://doi.org/10.1080/135062800394658

Simons, D. J., & Ambinder, M. S. (2005). Change blindness: Theory and consequences. *Current Directions in Psychological Science*, *14*(1), 44–48.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inatteintional blindness for dynamic events. *Perception*, *28*, 1059–1074. https://doi.org/10.1068/p2952

Simons, D. J., Chabris, C. F., Schnur, T., & Levin, D. T. (2002). Evidence for preserved representations in change blindness. *Consciousness and Cognition*, *11*(1), 78–97. https://doi.org/10.1006/ccog.2001.0533

Simons, D. J., Franconeri, S. L., & Reimer, R. L. (2000). Change blindness in the absence of a visual disruption. *Perception*, *29*(10), 1143–1154. https://doi.org/10.1068/p3104

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*(7), 261–267.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, *5*(4), 644–649. https://doi.org/10.3758/BF03208840

Simons, D. J., & Mitroff, S. R. (2001). The role of expectations in change detection and attentional capture. In M. Jenkin & L. Harris (Eds.), *Vision and attention* (pp. 189–207). New York: Springer. https://doi.org/https://doi.org/10.1007/978-0-387-21591-4

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16–20. https://doi.org/10.1016/j.tics.2004.11.006

Sinclair, L., & Kunda, Z. (1999). Reactions to a Black professional: Motivated inhibition and activation of conflicting stereotypes. *Journal of Personality and Social Psychology*, *77*(5), 885–904. https://doi.org/10.1037/0022-3514.77.5.885

Straube, S., & Fahle, M. (2011). Visual detection and identification are not the same: Evidence from psychophysics and fMRI. *Brain and Cognition*, *75*(1), 29–38. https://doi.org/10.1016/j.bandc.2010.10.004

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology*, *59*(11), 1931–1949. https://doi.org/10.1080/17470210500416342

Underwood, G., Foulsham, T., van Loon, E., Humphreys, L., & Bloyce, J. (2006). Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology*, *18*(3), 321–342. https://doi.org/10.1080/09541440500236661

Werner, S., & Thies, B. (2000). Is “Change Blindness” Attenuated by Domain-specific Expertise? An Expert-Novices Comparison of Change Detection in Football Images. *Visual Cognition*, *7*(1–3), 163–173. https://doi.org/10.1080/135062800394748

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, *81*(5), 815–827. https://doi.org/10.1037/0022-3514.81.5.815

Xu, Y. (2008). Distinctive neural mechanisms supporting visual object individuation and identification. *Journal of Cognitive Neuroscience*, *21*(3), 511–518. https://doi.org/10.1162/jocn.2008.21024

Zimmermann, E., Schnier, F., & Lappe, M. (2010). The contribution of scene context on change detection performance. *Vision Research, 50*(20), 2062–2068. https://doi.org/10.1016/j.visres.2010.07.019

Appendix B – Pre-test for used materials

Introduction

To test the effect of congruency between an object and the gender stereotype, it is necessary to ensure that the objects are in fact associated differently to either men or women – thus allowing their pairing with either gender and creating the congruency and incongruency conditions.

Gender roles contain behaviors and characteristics that are associated and expected from individuals according to their gender – expectations that are passed on between generations and are, consequently, culture-dependent (Neculăesei, 2015). These associations apply equally to objects of everyday use, with examples that populate all ages and contexts of our lives. For instance: it is expected of male children to play with toy cars, while female children are quickly associated with dolls; make-up is associated with women, while a tie is expected to be worn by a man.

Studies use a variety of methods to confirm that their stimuli are associated with a stereotype. Cheryan, Plaut, Davies, and Steele (2009, p. 1048), for example, asked a group of 33 students to list objects typically found in the dorm room or office of computer scientists, a science major, or a "computer science geek". Those most frequently mentioned were again rated by 20 students as to how typical of a computer science major they were, in a 7-item likert-type scale. Javadi and Wee (2012) asked participants to categorize images from a pre-existing catalog, shown in random order, as associated with the female or male gender. Similarly, Lemm and Dabady (2005) selected images from magazines and books and had them categorized as feminine, masculine, or gender neutral. Their 40 participants were instructed to judge the associations between object and gender as they "believe them to exist in this culture at this time".

Demonstrating this association between the stimuli and the categories in focus in each study is a vital step to assure task validity.

As such, this pre-test aims to confirm that a set of objects are strongly associated with each of the gender stereotypes (man and woman) in current Portuguese society.

Method

Participants

A sample of 33 participants was recruited by snowball sampling through e-mail and online social networks. Participants were of Portuguese nationality (23 female; 9 male, 1 undisclosed). Data on age is as follows: $\bar{x}$ = 35.55; SD = 13.39 (min 21; max 65).

Materials

Stimuli were either original photographs of objects, taken for this study, or images used with permission from Javadi and Wee (2012). The digital platform *Qualtrics* was used to create a survey adapted to both desktop and mobile devices.

Procedure

Participants received their participation link either by e-mail or through their social networking sites (e.g., Facebook). Upon clicking the link, the following instructions were shown:

Obrigado pelo seu interesse e participação.

Leia atentamente as seguintes instruções:
Neste estudo ser-lhe-á pedido que classifique um grupo de imagens, de acordo com a associação que acha que existe actualmente na sociedade.

Deverá categorizar as imagens de acordo com o género a que acha que mais se associa (Homem vs. Mulher);

A participação tem uma duração prevista de 3-4 minutos.

No final, ser-lhe-ão pedidos alguns dados demográficos (idade e género) não identificativos. Embora estes sejam facultativos, são relevantes para o estudo. A condifencialidade e anonimato dos dados é garantida.

Qualquer dúvida pode ser remetida para joaodavidmartins@gmail.com.

Pressione o botão abaixo para iniciar.

Figure B1. Instructions for the pre-test categorization task.

The next screen (Figure B2) listed the 40 objects to be categorized. Under each image, participants could choose one option between "Man" or "Woman". Each type of object was represented twice, to account for possible ambiguity (e.g., two different lipsticks were shown).

A final section collected information regarding participants' nationality, age, and gender.

Figure B2. Example of how the task looked when opened in mobile devices

Results

The following objects were categorized between 100% and 87,88% in the expected categories (Figures B3 and B4).

Male

Figure B3. Objects associated with "Man"

Female

Figure B4. Objects associated with "Woman"

Only one object was removed: a baby carriage was associated with the female gender by 71.43% of participants (28.57% associated it with "Man"). This was the only instance in which two versions of the same object differed in their categorization – a difference in frequency and not in gender, as both were still strongly associated with the "Woman" category. As the other exemplar of the same object scored higher (associated with the female gender by 88.57% of participants), and in order to avoid any confounding, that version was used in our task.

On average, both categories of objects average similarly in how frequently there were associated with the expected gender: male objects registered a 97,9% association the "Man" category; female objects registered a 98.2 association with the "Woman" category.

Discussion

All objects were categorized in accordance to expected associations, as well as confirming the results obtained by Javadi and Wee (2012) on the images made available by the authors.

Each object was represented by two exemplars. For each pair (aside from the removed object), there is no indication of ambiguity: both exemplars were always associated strongly with one gender exclusively.

Averages for each category (male and female objects) were calculated to establish equivalence of associations – in other words, to assure that, for instance, male objects were not associated with men at a higher magnitude than female objects were associated with women. Those averages indicated a balanced association between gender.

References

Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: how stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, *97*(6), 1045–1060. https://doi.org/10.1037/a0016239

Javadi, A. H., & Wee, N. (2012). Cross-category adaptation: Objects produce gender adaptation in the perception of faces. *PLoS ONE*, *7*(9), 3–10. https://doi.org/10.1371/journal.pone.0046079

Lemm, K. M., & Dabady, M. (2005). Gender picture priming: It works with denotative and connotative primes. *Social Cognition*, *23*(3), 218–241. https://doi.org/10.1521/soco.2005.23.3.218

Neculăesei, A.-N. (2015). Culture and Gender Role Differences. *Cross-Cultural Management Journal*, *17*(1), 31–35.

Appendix C – Images used (scenes)

The following pages list the images used in the task.

Each scene is represented in two rows, one per stereotype. Along each row, the first column shows the image without the changing object (i.e., the first to be shown in each trial). The second and third columns show, respectively, congruency (e.g., man with a male object) and incongruency (e.g., woman with male object).

Labels under each photo identify the scene (e.g., 01) and condition (e.g., MC for "man congruent"). For example: 05-WC refers to scene 5, figuring a woman with a congruent female object.


01-M

01-MC

01-MI

01-W

01-WC

01-WI

02-M

02-MC

02-MI

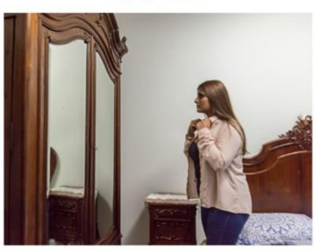
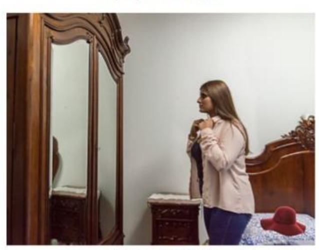
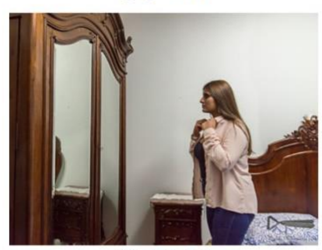
02-W

02-WC

02-WI

03-M

03-MC

03-MI

03-W

03-WC

03-WI


04-M

04-MC

04-MI

04-W

04-WC

04-WI

05-M

05-MC

05-MI

05-W

05-WC

05-WI

06-M

06-MC

06-MI

06-W

06-WC

06-WI

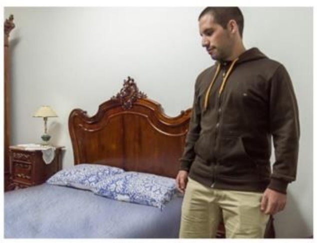
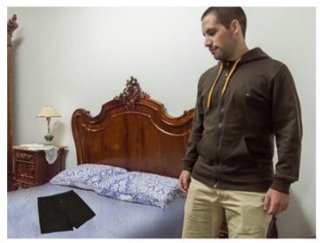
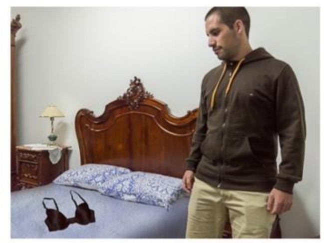

07-M

07-MC

07-MI

07-W

07-WC

07-WI

08-M

08-MC

08-MI

08-W

08-WC

08-WI

09-M

09-MC

09-MI

09-W

09-WC

09-WI


10-M

10-MC

10-MI

10-W

10-WC

10-WI

11-M

11-MC

11-MI

11-W

11-WC

11-WI

12-M

12-MC

12-MI

12-W

12-WC

12-WI

Appendix D – Test Instructions

Informed consent

Agradecemos a sua participação.

Esta sessão experimental é composta por 2 tarefas de percepção visual
e tem duração máxima de 30 minutos.

A sua participação é voluntária.
Os seus dados são anónimos e confidenciais.

Se concorda em prosseguir por favor prima "P".

Figure D1. Inform consent screen

Instructions

O objetivo desta tarefa é detetar mudança em imagens.

Cada imagem vai aparecer intermitentemente,
sempre no centro do ecrã.

Preste atenção ao conteúdo da imagem.

Quando achar que qualquer coisa mudou (mesmo sem saber o quê),
pressione a BARRA DE ESPAÇOS.

Tente ser o mais rápido(a) possível.

Figure D2. Instructions screen

Appendix E – Outputs from Statistical Analyses Performed

Output 1: Sample descriptive statistics

	Descriptive Statistics				
Variable	Valid N	Mean	Minimum	Maximum	Std.Dev.
Age	72	23.069	18	51	7.817

	Frequency table: gender		
Category	Count	% of all Cases	Cumulative % of All
Female	56	77.77778	77.7778
Male	16	22.22222	100.0000

Output 2: Power analysis to determine sample size (output from G*Power, Faul et al., 2007)

F tests – ANOVA: Repeated measures, within factors

Analysis:	A priori: Compute required sample size		
Input:	Effect size f(V)	=	0.3333333
	α err prob	=	0.05
	Power (1–β err prob)	=	0.8
	Number of groups	=	1
	Number of measurements	=	2
	Nonsphericity correction ϵ	=	1
Output:	Noncentrality parameter λ	=	8.1111095
	Critical F	=	3.9738970
	Numerator df	=	1.0000000
	Denominator df	=	72.0000000
	Total sample size	=	73

RT Analysis

Outputs 3 through 6 concern the section of the results titled "RT Analysis" (i.e., before exploring motivation and perceptive hypothesis, and before controlling for salience).

Output 3: RT Analysis – 2 (woman vs. man) x 2 (female vs. male) x 3 (400ms, 800ms, 1200ms) repeated measures ANOVA

	Repeated Measures Analysis of Variance with Effect Sizes and Powers Marked differenced are significant at p < .05							
Effect	SS	Degr. of Freedom	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0.05)
Intercept	1.1890E+09	1	1.1890E+09	1002.197	0.00000	0.9338	1002.197	1.0000
Error	8.4234E+07	71	1.1864E+06					
Object	4.2413E+06	1	4.2413E+06	10.699	0.00166	0.1310	10.699	0.8973
Error	2.8145E+07	71	3.9640E+05					
Threshold	2.4734E+07	2	1.2367E+07	24.431	0.00000	0.2560	48.861	1.0000
Error	7.1883E+07	142	5.0622E+05					
Stereotype	2.9972E+04	1	2.9972E+04	0.059	0.80936	0.0008	0.059	0.0566
Error	3.6290E+07	71	5.1113E+05					
Obj*Threshold	2.7114E+06	2	1.3557E+06	2.999	0.05300	0.0405	5.998	0.5743
Error	6.4192E+07	142	4.5205E+05					
Obj*Stereotype	1.2794E+05	1	1.2794E+05	0.519	0.47352	0.0073	0.519	0.1096
Error	1.7494E+07	71	2.4639E+05					
Threshold*Stereotype	2.9645E+06	2	1.4823E+06	3.743	0.02604	0.0501	7.487	0.6770
Error	5.6227E+07	142	3.9597E+05					
Obj*Threshold*Stereotype	1.3759E+04	2	6.8793E+03	0.016	0.98371	0.0002	0.033	0.0524
Error	5.9460E+07	142	4.1873E+05					

Output 4: RT Analysis – Fisher's LSD post-hoc test for Threshold

LSD test; variable: Threshold			
	400ms	800ms	1200ms
400ms		0.000000	0.000000
800ms	0.000000		0.470426
1200ms	0.000000	0.470426	

Output 5: RT Analysis – Fisher's LSD post-hoc test for Object gender * Threshold

	LSD test for Object gender * Threshold; variable RT							
Cell No.	Object gender	Threshold	{1} 816.39	{2} 1279.1	{3} 1213.6	{4} 1053.8	{5} 1262.3	{6} 1413.4
1	Female	400ms		0.0000	0.0000	0.0032	0.0000	0.0000
2	Female	800ms	0.0000		0.4093	0.0051	0.8321	0.0924
3	Female	1200ms	0.0000	0.4093		0.0456	0.5394	0.0128
4	Male	400ms	0.0032	0.0051	0.0456		0.0094	0.0000
5	Male	800ms	0.0000	0.8321	0.5394	0.0094		0.0586
6	Male	1200ms	0.0000	0.0924	0.0128	0.0000	0.0586	

Output 6: RT Analysis – Fisher's LSD post-hoc test for Threshold * Stereotype

	LSD test for Stereotype * Threshold; variable RT							
Cell No.	Threshold	Stereotype	{1} 886.95	{2} 983.24	{3} 1347.7	{4} 1193.8	{5} 1267.0	{6} 1359.9
1	400ms	Woman		0.1962	0.0000	0.0001	0.0000	0.0000
2	400ms	Man	0.1962		0.0000	0.0052	0.0002	0.0000
3	800ms	Woman	0.0000	0.0000		0.0398	0.2786	0.8688
4	800ms	Man	0.0001	0.0052	0.0398		0.3251	0.0266
5	1200ms	Woman	0.0000	0.0002	0.2786	0.3251		0.2122
6	1200ms	Man	0.0000	0.0000	0.8688	0.0266	0.2122	

Motivational hypothesis

The outputs below, from 7 through 12, correspond to the section in the Results titled "Motivational hypothesis". Each output is identified as referring to the female or male sample.

Output 7: Female sample – 2 (woman vs. man) x 2 (female vs. male) x 3 (400ms, 800ms, 1200ms) repeated measures ANOVA

Effect	Repeated Measures Analysis of Variance with Effect Sizes and Powers Marked differences are significant at p < .05							
	SS	Degr. of Freedom	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0.05)
Intercept	865765675	1	865765675	776.3774	0.0000	0.93497	776.3774	1.0000
Error	60217296	54	1115135					
Object	4638177	1	4638177	12.1949	0.0010	0.18423	12.1949	0.9291
Error	20538190	54	380337					
Threshold	16969349	2	8484674	19.3183	0.0000	0.26349	38.6367	0.9999
Error	47433949	108	439203					
Stereotype	35651	1	35651	0.0843	0.7727	0.00156	0.0843	0.0594
Error	22837729	54	422921					
Object*Threshold	1993315	2	996657	2.2666	0.1086	0.04028	4.5333	0.4524
Error	47488324	108	439707					
Object*Stereotye	93457	1	93457	0.3631	0.5493	0.00668	0.3631	0.0910
Error	13898905	54	257387					
Threshold*Stereotype	1451347	2	725673	1.8830	0.1571	0.03370	3.7660	0.3843
Error	41620969	108	385379					
Obj*Threshold*Stereotype	58485	2	29242	0.0668	0.9354	0.00124	0.1336	0.0599
Error	47262782	108	437618					

Output 8: Female sample – Means for Object gender

Cell No.	Object gender; Unweighted Means					
	Object gender	RT Mean	RT Std.Err.	RT -95.00%	RT +95.00%	N
1	Female	1061.493	40.6702	979.954	1143.032	55
2	Male	1229.154	53.6439	1121.604	1336.704	55

Output 9: Female sample – Means for Threshold

Cell No.	Threshold; Unweighted Means					
	Threshold	RT Mean	RT Std.Err.	RT -95.00%	RT +95.00%	N
1	400ms	922.005	44.9052	831.975	1012.034	55
2	800ms	1291.090	61.0708	1168.650	1413.529	55
3	1200ms	1222.876	57.5798	1107.436	1338.317	55

Output 10: Female sample – Fisher's LSD post-hoc test for Object gender * Threshold

Cell No.	LSD test for Object gender * Threshold; variable RT							
	Object gender	Threshold	{1} 782.56	{2} 1282.1	{3} 1119.8	{4} 1061.5	{5} 1300.1	{6} 1325.9
1	Female	400ms		0.000000	0.000265	0.002325	0.000000	0.000000
2	Female	800ms	0.000000		0.072380	0.015176	0.840680	0.625004
3	Female	1200ms	0.000265	0.072380		0.515130	0.046284	0.023098
4	Male	400ms	0.002325	0.015176	0.515130		0.008781	0.003808
5	Male	800ms	0.000000	0.840680	0.046284	0.008781		0.773389
6	Male	1200ms	0.000000	0.625004	0.023098	0.003808	0.773389	

Output 11: Male sample – 2 (woman vs. man) x 2 (female vs. male) x 3 (400ms, 800ms, 1200ms) repeated measures ANOVA

Effect	Repeated Measures Analysis of Variance with Effect Sizes and Powers Marked differences are significant at p < .05							
	SS	Degr. of Freedom	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0.05)
Intercept	316928554	1	316928554	223.2773	0.000000	0.937048	223.2773	1.000000
Error	21291584	15	1419439					
Object gender	72911	1	72911	0.1562	0.698270	0.010304	0.1562	0.065871
Error	7002966	15	466864					
Threshold	18206771	2	9103386	15.3896	0.000025	0.506410	30.7792	0.998440
Error	17745841	30	591528					
Stereotupe	217891	1	217891	0.2533	0.622044	0.016609	0.2533	0.075876
Error	12901153	15	860077					
Object*Threshold	2800889	2	1400445	1.8459	0.175367	0.109577	3.6918	0.353973
Error	22760102	30	758670					
Object*Stereotype	93972	1	93972	0.3629	0.555910	0.023620	0.3629	0.087260
Error	3884450	15	258963					
Threshold*Stereotype	1924665	2	962332	2.2069	0.127590	0.128258	4.4139	0.414813
Error	13081486	30	436050					
Obj*Threshold*Stereotype	278384	2	139192	0.3656	0.696846	0.023792	0.7311	0.103260
Error	11422467	30	380749					

Output 12: Male sample – Fisher's LSD post-hoc test for Threshold

Cell No.	LSD test for Threshold; variable RT			
	Threshold	{1} 943.20	{2} 1221.6	{3} 1689.5
1	400ms		0.049415	0.000006
2	800ms	0.049415		0.001725
3	1200ms	0.000006	0.001725	

Perceptive hypothesis

The outputs below, from 13 through 15, correspond to the section in the Results titled "Perceptive hypothesis".

Output 13: Perceptive hypothesis – Results of repeated measures ANOVAS: 12 (scenes) x 2 (Stereotype), and 12 (scenes) x 2 (Object gender).

	Repeated Measures Analysis of Variance with Effect Sizes and Powers Marked differences are significant at p < .05							
Effect	SS	Degr. of Freedom	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0.05)
Intercept	2.37191E+09	1.0	2.3719E+09	1038.110	0.000000	0.93937	1038.110	1.0000
Error	1.53084E+08	67.0	2.2848E+06					
Scene	3.51105E+08	11.0	3.1919E+07	31.344	0.000000	0.31872	344.780	1.0000
Error	7.50520E+08	737.0	1.0183E+06					
Stereotype	4.75088E+03	1.0	4.7509E+03	0.005	0.945508	0.00007	0.005	0.0505
Error	6.76279E+07	67.0	1.0094E+06					
Scene*Stereotype	2.73486E+07	11.0	2.4862E+06	3.203	0.000285	0.04563	35.236	0.9927
Error	5.72026E+08	737.0	7.7615E+05					
Object	6.80141E+06	1.0	6.8014E+06	8.1679	0.005591	0.10317	8.1679	0.8049
Error	5.91214E+07	71.0	8.3270E+05					
Scene*Object	3.88312E+07	11.0	3.5301E+06	4.3406	0.000003	0.05761	47.7467	0.9996
Error	6.35168E+08	781.0	8.1328E+05					

Output 14: Perceptive hypothesis – Fisher's LSD post-hoc test for Scene * Object gender

LSD test for Scene*Object gender; variable RT

	Scene	Object gender	{1} 1151.6	{2} 1253.3	{3} 879.87	{4} 1212.6	{5} 625.28	{6} 1342.2	{7} 980.67	{8} 1015.4	{9} 1434.5	{10} 1198.1	{11} 1429.9	{12} 1371.7	{13} 616.53	{14} 1331.1	{15} 572.07	{16} 586.19	{17} 1701.3	{18} 1801.5	{19} 497.59	{20} 576.89	{21} 2171.1	{22} 1994.3	{23} 983.07	{24} 865.96
1	1	Female		0.4988	0.0710	0.6851	0.0005	0.2051	0.2558	0.3653	0.0602	0.7572	0.0645	0.1435	0.0004	0.2327	0.0001	0.0002	0.0003	0.0000	0.0000	0.0001	0.0000	0.0000	0.2626	0.0578
2	1	Male	0.4988		0.0132	0.7864	0.0000	0.5543	0.0701	0.1139	0.2284	0.7134	0.2405	0.4312	0.0000	0.6048	0.0000	0.0000	0.0030	0.0003	0.0000	0.0000	0.0000	0.0000	0.0726	0.0101
3	2	Female	0.0710	0.0132		0.0272	0.0907	0.0022	0.5027	0.3674	0.0002	0.0346	0.0003	0.0011	0.0802	0.0028	0.0409	0.0511	0.0000	0.0000	0.0112	0.0442	0.0000	0.0000	0.4926	0.9263
4	2	Male	0.6851	0.7864	0.0272		0.0001	0.3886	0.1233	0.1901	0.1402	0.9232	0.1487	0.2901	0.0001	0.4305	0.0000	0.0000	0.0012	0.0001	0.0000	0.0000	0.0000	0.0000	0.1272	0.0214
5	3	Female	0.0005	0.0000	0.0907	0.0001		0.0000	0.0183	0.0096	0.0000	0.0001	0.0000	0.0000	0.9536	0.0000	0.7234	0.7949	0.0000	0.0000	0.3958	0.7475	0.0000	0.0000	0.0175	0.1097
6	3	Male	0.2051	0.5543	0.0022	0.3886	0.0000		0.0164	0.0300	0.5396	0.3378	0.5600	0.8447	0.0000	0.9411	0.0000	0.0000	0.0171	0.0023	0.0000	0.0000	0.0000	0.0000	0.0171	0.0016
7	4	Female	0.2558	0.0701	0.5027	0.1233	0.0183	0.0164		0.8172	0.0026	0.1485	0.0029	0.0095	0.0156	0.0200	0.0067	0.0088	0.0000	0.0000	0.0014	0.0074	0.0000	0.0000	0.9873	0.4456
8	4	Male	0.3653	0.1139	0.3674	0.1901	0.0096	0.0300	0.8172		0.0054	0.2247	0.0060	0.0180	0.0081	0.0360	0.0033	0.0044	0.0000	0.0000	0.0006	0.0036	0.0000	0.0000	0.8296	0.3203
9	5	Female	0.0602	0.2284	0.0002	0.1402	0.0000	0.5396	0.0026	0.0054		0.1162	0.9755	0.6762	0.0000	0.4918	0.0000	0.0000	0.0762	0.0148	0.0000	0.0000	0.0000	0.0002	0.0028	0.0002
10	5	Male	0.7572	0.7134	0.0346	0.9232	0.0001	0.3378	0.1485	0.2247	0.1162		0.1234	0.2485	0.0001	0.3764	0.0000	0.0001	0.0009	0.0001	0.0000	0.0000	0.0000	0.0000	0.1530	0.0274
11	6	Female	0.0645	0.2405	0.0003	0.1487	0.0000	0.5600	0.0029	0.0060	0.9755	0.1234		0.6987	0.0000	0.5113	0.0000	0.0000	0.0713	0.0136	0.0000	0.0000	0.0000	0.0002	0.0030	0.0002
12	6	Male	0.1435	0.4312	0.0011	0.2901	0.0000	0.8447	0.0095	0.0180	0.6762	0.2485	0.6987		0.0000	0.7873	0.0000	0.0000	0.0286	0.0044	0.0000	0.0000	0.0000	0.0000	0.0099	0.0008
13	7	Female	0.0004	0.0000	0.0802	0.0001	0.9536	0.0000	0.0156	0.0081	0.0000	0.0001	0.0000	0.0000		0.0000	0.7674	0.8401	0.0000	0.0000	0.4290	0.7920	0.0000	0.0000	0.0150	0.0974
14	7	Male	0.2327	0.6048	0.0028	0.4305	0.0000	0.9411	0.0200	0.0360	0.4918	0.3764	0.5113	0.7873	0.0000		0.0000	0.0000	0.0140	0.0018	0.0000	0.0000	0.0000	0.0000	0.0208	0.0020
15	8	Female	0.0001	0.0000	0.0409	0.0000	0.7234	0.0000	0.0067	0.0033	0.0000	0.0000	0.0000	0.0000	0.7674	0.0000		0.9252	0.0000	0.0000	0.6204	0.9744	0.0000	0.0000	0.0064	0.0509
16	8	Male	0.0002	0.0000	0.0511	0.0000	0.7949	0.0000	0.0088	0.0044	0.0000	0.0001	0.0000	0.0000	0.8401	0.0000	0.9252		0.0000	0.0000	0.5557	0.9506	0.0000	0.0000	0.0084	0.0631
17	9	Female	0.0003	0.0030	0.0000	0.0012	0.0000	0.0171	0.0000	0.0000	0.0762	0.0009	0.0713	0.0286	0.0000	0.0140	0.0000	0.0000		0.5055	0.0000	0.0000	0.0018	0.0516	0.0000	0.0000
18	9	Male	0.0000	0.0003	0.0000	0.0001	0.0000	0.0023	0.0000	0.0000	0.0148	0.0001	0.0136	0.0044	0.0000	0.0018	0.0000	0.0000	0.5055		0.0000	0.0000	0.0141	0.1998	0.0000	0.0000
19	10	Female	0.0000	0.0000	0.0112	0.0000	0.3958	0.0000	0.0014	0.0006	0.0000	0.0000	0.0000	0.0000	0.4290	0.0000	0.6204	0.5557	0.0000	0.0000		0.5979	0.0000	0.0000	0.0013	0.0145
20	10	Male	0.0001	0.0000	0.0442	0.0000	0.7475	0.0000	0.0074	0.0036	0.0000	0.0000	0.0000	0.0000	0.7920	0.0000	0.9744	0.9506	0.0000	0.0000	0.5979		0.0000	0.0000	0.0070	0.0548
21	11	Female	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0018	0.0141	0.0000	0.0000		0.2399	0.0000	0.0000
22	11	Male	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0516	0.1998	0.0000	0.0000	0.2399		0.0000	0.0000
23	12	Female	0.2626	0.0726	0.4926	0.1272	0.0175	0.0171	0.9873	0.8296	0.0028	0.1530	0.0030	0.0099	0.0150	0.0208	0.0064	0.0084	0.0000	0.0000	0.0013	0.0070	0.0000	0.0000		0.4361
24	12	Male	0.0578	0.0101	0.9263	0.0214	0.1097	0.0016	0.4456	0.3203	0.0002	0.0274	0.0002	0.0008	0.0974	0.0020	0.0509	0.0631	0.0000	0.0000	0.0145	0.0548	0.0000	0.0000	0.4361	

Output 15: Perceptive hypothesis – Fisher's LSD post-hoc test for Scene * Stereotype

LSD test for Scene * Stereotype; variable RT

	Scene	Stereotype	{1} 1130.2	{2} 1351.3	{3} 1135.7	{4} 1016.1	{5} 832.10	{6} 1208.0	{7} 920.65	{8} 1129.2	{9} 1292.9	{10} 1441.3	{11} 1541.8	{12} 1493.9	{13} 1015.7	{14} 1008.8	{15} 612.47	{16} 587.35	{17} 1828.4	{18} 1964.5	{19} 570.31	{20} 540.06	{21} 2560.3	{22} 1853.6	{23} 1005.7	{24} 893.16
1	1	Female		0.1439	0.9712	0.4501	0.0488	0.6068	0.1658	0.9947	0.2820	0.0399	0.0066	0.0163	0.4486	0.4217	0.0006	0.0003	0.0000	0.0000	0.0002	0.0001	0.0000	0.0000	0.4102	0.1171
2	1	Male	0.1439		0.1540	0.0268	0.0006	0.3434	0.0045	0.1421	0.6993	0.5515	0.2077	0.3455	0.0266	0.0237	0.0000	0.0000	0.0017	0.0001	0.0000	0.0000	0.0000	0.0009	0.0225	0.0025
3	2	Female	0.9712	0.1540		0.4288	0.0449	0.6322	0.1551	0.9659	0.2984	0.0435	0.0074	0.0180	0.4273	0.4012	0.0006	0.0003	0.0000	0.0000	0.0002	0.0001	0.0000	0.0000	0.3900	0.1089
4	2	Male	0.4501	0.0268	0.4288		0.2238	0.2043	0.5280	0.4541	0.0673	0.0050	0.0005	0.0016	0.9980	0.9616	0.0077	0.0047	0.0000	0.0000	0.0033	0.0017	0.0000	0.0000	0.9455	0.4163
5	3	Female	0.0488	0.0006	0.0449	0.2238		0.0131	0.5580	0.0496	0.0024	0.0001	0.0000	0.0000	0.2248	0.2427	0.1465	0.1057	0.0000	0.0000	0.0836	0.0536	0.0000	0.0000	0.2509	0.6862
6	3	Male	0.6068	0.3434	0.6322	0.2043	0.0131		0.0576	0.6022	0.5745	0.1230	0.0275	0.0589	0.2034	0.1877	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1810	0.0375
7	4	Female	0.1658	0.0045	0.1551	0.5280	0.5580	0.0576		0.1679	0.0140	0.0006	0.0000	0.0002	0.5296	0.5599	0.0417	0.0277	0.0000	0.0000	0.0207	0.0120	0.0000	0.0000	0.5736	0.8557
8	4	Male	0.9947	0.1421	0.9659	0.4541	0.0496	0.6022	0.1679		0.2790	0.0392	0.0065	0.0160	0.4526	0.4256	0.0007	0.0004	0.0000	0.0000	0.0002	0.0001	0.0000	0.0000	0.4140	0.1186
9	5	Female	0.2820	0.6993	0.2984	0.0673	0.0024	0.5745	0.0140	0.2790		0.3263	0.0999	0.1838	0.0669	0.0604	0.0000	0.0000	0.0004	0.0000	0.0000	0.0000	0.0000	0.0002	0.0577	0.0083
10	5	Male	0.0399	0.5515	0.0435	0.0050	0.0001	0.1230	0.0006	0.0392	0.3263		0.5061	0.7278	0.0050	0.0043	0.0000	0.0000	0.0106	0.0006	0.0000	0.0000	0.0000	0.0065	0.0041	0.0003
11	6	Female	0.0066	0.2077	0.0074	0.0005	0.0000	0.0275	0.0000	0.0065	0.0999	0.5061		0.7513	0.0005	0.0004	0.0000	0.0000	0.0582	0.0053	0.0000	0.0000	0.0000	0.0394	0.0004	0.0000
12	6	Male	0.0163	0.3455	0.0180	0.0016	0.0000	0.0589	0.0002	0.0160	0.1838	0.7278	0.7513		0.0016	0.0014	0.0000	0.0000	0.0271	0.0019	0.0000	0.0000	0.0000	0.0175	0.0013	0.0001
13	7	Female	0.4486	0.0266	0.4273	0.9980	0.2248	0.2034	0.5296	0.4526	0.0669	0.0050	0.0005	0.0016		0.9636	0.0078	0.0047	0.0000	0.0000	0.0033	0.0017	0.0000	0.0000	0.9475	0.4177
14	7	Male	0.4217	0.0237	0.4012	0.9616	0.2427	0.1877	0.5599	0.4256	0.0604	0.0043	0.0004	0.0014	0.9636		0.0089	0.0054	0.0000	0.0000	0.0038	0.0020	0.0000	0.0000	0.9839	0.4444
15	8	Female	0.0006	0.0000	0.0006	0.0077	0.1465	0.0001	0.0417	0.0007	0.0000	0.0000	0.0000	0.0000	0.0078	0.0089		0.8680	0.0000	0.0000	0.7803	0.6319	0.0000	0.0000	0.0094	0.0636
16	8	Male	0.0003	0.0000	0.0003	0.0047	0.1057	0.0000	0.0277	0.0004	0.0000	0.0000	0.0000	0.0000	0.0047	0.0054	0.8680		0.0000	0.0000	0.9102	0.7544	0.0000	0.0000	0.0058	0.0433
17	9	Female	0.0000	0.0017	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0106	0.0582	0.0271	0.0000	0.0000	0.0000	0.0000		0.3680	0.0000	0.0000	0.0000	0.8674	0.0000	0.0000
18	9	Male	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0053	0.0019	0.0000	0.0000	0.0000	0.0000	0.3680		0.0000	0.0000	0.0001	0.4634	0.0000	0.0000
19	10	Female	0.0002	0.0000	0.0002	0.0033	0.0836	0.0000	0.0207	0.0002	0.0000	0.0000	0.0000	0.0000	0.0033	0.0038	0.7803	0.9102	0.0000	0.0000		0.8414	0.0000	0.0000	0.0041	0.0329
20	10	Male	0.0001	0.0000	0.0001	0.0017	0.0536	0.0000	0.0120	0.0001	0.0000	0.0000	0.0000	0.0000	0.0017	0.0020	0.6319	0.7544	0.0000	0.0000	0.8414		0.0000	0.0000	0.0021	0.0197
21	11	Female	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000		0.0000	0.0000	0.0000
22	11	Male	0.0000	0.0009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0065	0.0394	0.0175	0.0000	0.0000	0.0000	0.0000	0.8674	0.4634	0.0000	0.0000	0.0000		0.0000	0.0000
23	12	Female	0.4102	0.0225	0.3900	0.9455	0.2509	0.1810	0.5736	0.4140	0.0577	0.0041	0.0004	0.0013	0.9475	0.9839	0.0094	0.0058	0.0000	0.0000	0.0041	0.0021	0.0000	0.0000		0.4565
24	12	Male	0.1171	0.0025	0.1089	0.4163	0.6862	0.0375	0.8557	0.1186	0.0083	0.0003	0.0000	0.0001	0.4177	0.4444	0.0636	0.0433	0.0000	0.0000	0.0329	0.0197	0.0000	0.0000	0.4565	

Re-testing the hypothesis

The outputs below, from 16 and 17, correspond to the section in the Results titled "Re-testing the hypothesis controlling for salience differences". These analyses were conducted after removal of scenes 2, 3, 7, and 11 – as these were found to contain interfering saliency issues.

Output 16: Re-test – 2 (woman vs. man) x 2 (female vs. male) x 3 (400ms, 800ms, 1200ms) repeated measures ANOVA

	Repeated Measures Analysis of Variance with Effect Sizes and Powers Marked differences are significant at p < .05							
Effect	SS	Degr. of Freedom	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0.05)
Intercept	1.0366E+09	1	1.0366E+09	1008.665	0.00000	0.9360	1008.665	1.0000
Error	7.0909E+07	69	1.0277E+06					
Stereotype	7.4997E+05	1	7.4997E+05	2.139	0.14813	0.0301	2.139	0.3026
Error	2.4192E+07	69	3.5061E+05					
Object gender	7.0084E+03	1	7.0084E+03	0.015	0.90299	0.0002	0.015	0.0517
Error	3.2312E+07	69	4.6829E+05					
Threshold	4.0524E+07	2	2.0262E+07	37.633	0.00000	0.3529	75.266	1.0000
Error	7.4300E+07	138	5.3841E+05					
Stereotype*Object	1.5406E+04	1	1.5406E+04	0.037	0.84853	0.0005	0.037	0.0541
Error	2.8923E+07	69	4.1918E+05					
Stereotype*Threshold	7.9010E+05	2	3.9505E+05	0.861	0.42502	0.0123	1.722	0.1957
Error	6.3323E+07	138	4.5886E+05					
Object*Threshold	1.5812E+06	2	7.9062E+05	1.956	0.14530	0.0276	3.912	0.3996
Error	5.5775E+07	138	4.0417E+05					
Stereotype*Obj*Threshold	5.2278E+05	2	2.6139E+05	0.717	0.49010	0.0103	1.434	0.1693
Error	5.0320E+07	138	3.6463E+05					

Output 17: Re-test – Fisher's LSD post-hoc test for Threshold

	LSD test for Threshold; variable RT Marked differences are significant at p < .05			
Cell No.	Threshold	{1} 929.23	{2} 983.46	{3} 1419.9
1	400ms		0.3834	0.0000
2	800ms	0.3834		0.0000
3	1200ms	0.0000	0.0000	

Output 18: Pearson's correlation analysis (between used Threshold and observed RTs) for all 8

studies (23 observations).

Variable	Correlations Marked correlations are significant at p < .05 N=23 (Casewise deletion of missing data)			
	Means	Std.Dev.	Threshold	RTs
Threshold	618.261	340.356	1.0000	0.8600
RTs	5541.087	4112.203	0.8600	1.0000

Output 19: Pearson's correlation analysis (between Threshold and observed RTs) after removing observations from Vierck and Kiesel (2008).

Variable	Correlations Marked correlations are significant at p < .05000 N=17 (Casewise deletion of missing data)			
	Means	Std.Dev.	Threshold	RTs
Threshold	497.647	102.135	1.0000	0.3278
RTs	3442.294	1174.865	0.3278	1.0000