



The woman who wasn't there: Converging evidence that subliminal social comparison affects self-evaluation[☆]



Armand Chatard^{a,*}, Yvana Bocage-Barthélémy^a, Leila Selimbegović^a, Serge Guimond^b

^a Université de Poitiers, Université de Tours, & CNRS, MSHS, Bat A5, 5 rue Théodore Lefebvre, TSA 21103, 86073 Poitiers Cedex 9, France

^b Université Clermont Auvergne & CNRS, 34 rue avenue Carnot, 63037 Clermont-Ferrand, France

ARTICLE INFO

Keywords:

Subliminal social comparison
Automatic processes
Explicit self-evaluations
Thin-ideal
Body appearance anxiety
Replication crisis

ABSTRACT

Although social comparison is often considered as an automatic process, the evidence in support of this idea is weak and inconclusive. In this paper, we reexamined the question of automaticity in social comparison by testing the hypothesis that subliminal social comparison affects explicit self-evaluations. In two high-powered experiments, young women were subliminally exposed (or not) to a high standard of comparison (media images of ultra-thin women). Next, they made explicit self-evaluations of their body appearance anxiety. Using both between-participants (Experiment 1) and within-participant (Experiment 2) designs, we found converging evidence that subliminal exposure to the thin ideal increases body appearance anxiety in women. Using Bayes factors as measures of evidence, the present experiments provided substantial (Experiment 1) and very strong (Experiment 2) evidence that social comparison takes place outside awareness and affects explicit self-evaluations. The present experiments can be easily replicated using a standardized procedure (replication script) that is publicly available on the Open Science Framework. We discuss how these findings contribute to reestablish confidence in the modern view of social comparison as an automatic process.

1. Introduction

People often evaluate themselves by comparing their own attributes and abilities with those of others, a process called social comparison (for comprehensive reviews, see Guimond, 2006; Mussweiler, 2003; Suls, Martin, & Wheeler, 2002). One of the most intriguing hypotheses of modern research on social comparison is the idea that it can operate automatically and without awareness (Alicke, 2007). If social comparison is automatic, then it is likely to operate spontaneously (without intention), unconsciously, effortlessly, and in an uncontrollable manner (Bargh, 1994). In the early 2000s, the notion of automatic social comparison was widely accepted, and it still remains quite popular today (see Want, 2009). However, a number of research findings previously considered as robust are being questioned in the framework of the replication crisis that the field of social psychology is currently undergoing (Open Science Collaboration, 2015). Indeed, several papers that purportedly demonstrated the automaticity of social comparison were retracted following the revelation of Diederick Stapel's massive fraud (Callaway, 2011). In addition, other, trustworthy experiments

relevant to this issue have been conducted at a time when methodological standards were much less stringent than today. As a consequence, even if they provide *prima facie* evidence in favor of the automaticity of social comparison, this evidence can be questioned on the basis of, for instance, power analyses. In this context, we revisit the hypothesis that social comparison is automatic from the lens of a skeptical researcher. After reviewing the available empirical reports, we conclude that there is at present no solid evidence to substantiate the claim that social comparison is automatic. Then, we report the results of two high-powered experiments conducted in an effort to provide further evidence in favor of or against the hypothesis that social comparison is an automatic process.

1.1. Classical and modern views of social comparison

Social comparison has not always been considered as an automatic process. Festinger (1954), the forefather of social comparison theory, conceptualized social comparison as a deliberate and controlled process. In this perspective (hereafter referred to as the classical perspective),

[☆] All materials, data, and preregistrations are available at <https://osf.io/gjc5b/>. This research was supported by a grant from the Agence Nationale de la Recherche (BIPA project, no ANR-13-BSH2-0008). We thank Amandine Gaboriau, Léa Gaillard, Emilie Juteau, Florie Bordier, Yves Almecija, and Marie-Françoise Crété for help with data collection and the preparation of this manuscript.

* Corresponding author.

E-mail addresses: armand.chatard@univ-poitiers.fr (A. Chatard), yvana.bocage.barthelemy@univ-poitiers.fr (Y. Bocage-Barthélémy), leila.selimbegovic@univ-poitiers.fr (L. Selimbegović), serge.guimond@uca.fr (S. Guimond).

<http://dx.doi.org/10.1016/j.jesp.2017.05.005>

Received 15 December 2016; Received in revised form 10 May 2017; Accepted 10 May 2017

Available online 15 June 2017

0022-1031/ © 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

individuals do not automatically compare themselves with others. Rather, they intentionally select comparisons with similar others and neglect comparisons with dissimilar targets of comparison, because such comparisons are deemed as non-diagnostic for self-evaluation. Several effects documented in this line of research are generally considered robust (or highly replicable), such as effects of upward and downward social comparisons (Gibbons & Gerrard, 1989; Morse & Gergen, 1970; Taylor & Lobel, 1989) or effects of social comparisons with ingroup and outgroup members (Guimond, Chatard, Martinot, Crisp, & Redersdorff, 2006; Mussweiler & Bodenhausen, 2002).

In contrast to the classical perspective, Gilbert, Giesler, and Morris (1995) have advanced the idea that social comparison is a relatively spontaneous, effortless, and unintentional process that may occur even when the comparison is deemed as irrelevant or nondiagnostic. In this perspective (hereafter referred to as the modern perspective), social comparison involves two different and consecutive processes. The first one is automatic. When exposed to another person, or another group of persons, individuals automatically compare themselves to others. The second, correction process, is controlled and occurs only after initial comparisons are made. If and when individuals realize that the comparison is not relevant or nondiagnostic for self-evaluation, they make a correction to their self-evaluation to take into account the nondiagnostic feature of the comparison. In other words, they intentionally “undo” the comparison after it has been made.

To test this reasoning, Gilbert et al. (1995) conducted two experiments in which they manipulated cognitive load (low or high) in between-participants designs. High cognitive load was expected to disrupt controlled processes, and thus the correction made after initial social comparisons with nondiagnostic targets of comparison. When participants were able to make a correction (under low cognitive load), their self-evaluation was unaffected by nondiagnostic comparison information. However, when participants' ability to make a correction was reduced (under high cognitive load), their self-evaluation was affected by the social comparison with another person, even when the comparison was nondiagnostic. In other words, the social comparison process appeared to be efficient (resource-independent) and uncontrollable (occurred even if explicitly irrelevant). These findings provided evidence in favor of the modern view of social comparison as an automatic process.

1.2. A critical look at the automaticity of social comparison

Perhaps because it seems intuitively appealing, the modern perspective of social comparison has been quite popular in the past. However, the automaticity of social comparison is often assumed but rarely demonstrated (see Want, Botres, Vahedi, & Middleton, 2015). Bargh (1994) defined automaticity through 4 specific criteria: the absence of awareness, efficiency, unintentionality, and uncontrollability. Admittedly, a given process can have some features of automaticity and not others (Bargh, 1989). These four features are clearly distinct at a conceptual level. However, when a person is unaware of the presence of a stimulus, s/he is unlikely to be able to use it intentionally, or to attempt to control its influence on his or her cognition, emotion, or behavior (Bargh, 1994). This, of course, does not imply that such an influence cannot be controlled when the person is aware of the stimulus. Perhaps for this reason, previous research has typically focused on two criteria of automaticity: awareness and efficiency. By definition, automatic mental processes require relatively few mental resources (i.e., they are not impaired by cognitive load), and operate without awareness (or subliminally). Thus, social comparison can be considered as an automatic process if it can be shown that it requires very few mental resources, or that it takes place without awareness – upon subliminal presentation of stimuli (bearing in mind that absence of awareness makes intention and control highly unlikely).

Research examining whether social comparison is an efficient automatic process has produced inconsistent results. Recently, Want

and colleagues found no evidence that social comparisons are efficient mental processes (Want & Saiphoo, 2017; Want et al., 2015). In these high-powered experiments, exposure to a high standard of comparison (media images of ultra-thin women) decreased explicit self-evaluation and impaired mood among women who were not cognitively busy (low cognitive load), but not among those who were (high cognitive load). The results of these experiments cast doubt on the idea that social comparison is automatic.

Recently, we conducted a similar experiment with an implicit rather than an explicit measure as the dependent variable (Bocage-Barthélémy et al., 2017). In this high-powered experiment, a lexical decision task was used to assess the cognitive accessibility of negative words. Exposure to the high standard of comparison (the thin ideal), compared to a no comparison condition (pictures of women's fashion accessories), caused greater accessibility of negative words, but only among women who were cognitively busy (high cognitive load). There was no effect of the comparison among those who were not cognitively busy (low cognitive load). Here, the findings were more consistent with the modern view than with the classical view of social comparison.

All in all, these conflicting findings are puzzling. It remains unclear why Want et al. (2015) and Want and Saiphoo (2017) found no effects of social comparison under high cognitive load, while Bocage-Barthélémy et al. (2017) found no effects of social comparison under low cognitive load. Differences between procedures and dependent variables may account for this discrepancy. Clearly, further research is needed to address this issue. At present, the only conclusion that can be drawn is that research examining whether social comparison is an efficient process has produced inconsistent findings.

Using subliminal presentation of comparison information, a number of experiments have tested whether social comparison operates without awareness (Blanton & Stapel, 2008; Jansen & de Vries, 2002; Mussweiler, Rüter, & Epstude, 2004; Stapel & Blanton, 2004). However, some experiments published by Stapel and his coauthors have been retracted after he admitted fabricating his data. In addition, the few remaining experiments are suboptimal because they are underpowered. They all included about 20 participants per condition. Although it was a common research practice at the time these experiments were conducted, such low-powered experiments are clearly problematic. Experiments with such small sample sizes have only 23% power to detect the typical small-to-medium effect size found in social psychology (Richard, Bond, & Stokes-Zoota, 2003). Of course, with only 23% statistical power, most experiments would fail most of the time. Many failed experiments may not have been reported, making it difficult to gauge the true effect size from the published literature, or to draw any firm conclusions about the awareness of social comparison processes from this literature.

Jansen and de Vries (2002) conducted a single experiment with a total of 59 participants, randomly assigned into one of three experimental conditions. They found no effect of subliminal social comparisons and concluded that subliminal social comparisons do not affect self-evaluation. It is quite possible that the true effect size is not different from zero. However, an alternative explanation is that these researchers did not have enough power to detect a significant effect in their experiment. Indeed, the only conclusion that can be drawn from this experiment is that subliminal social comparisons do not have an extraordinarily large effect on self-evaluation.

At the same time, Mussweiler et al. (2004) reported different results. Their sample size was even smaller, yet they reported 3 successful experiments. Mussweiler et al. predicted that subliminal social comparisons with moderately low or high standards would lead to assimilation effects (participants evaluate themselves in line with the activated standard), whereas subliminal social comparisons with extremely low or high standards would lead to contrast effects (self-evaluations are contrasted away from the activated standards). In a first experiment (N = 32) including two independent conditions, they found that German students primed with the name of a moderately high standard of aggressiveness (the

German TV-detective Schimanski) evaluated themselves as being more aggressive than those primed with the name of a moderately low standard of aggressiveness (the German pop-singer Guiseppe Horn). In a second experiment ($N = 55$) including 4 independent conditions, they also found assimilation effects when the standard was moderately low (Bill Clinton) or high (the former race car driver Nicki Lauda) on the dimension of athletics abilities. However, they found contrast effects when the standards were extremely low (Pope John Paul II) or high (Michael Jordan) on the dimension of athletic ability. In the third experiment ($N = 47$) including 4 independent conditions, they found assimilation effects when the standard was moderately low (the German pop-singer Nena) or high (Arnold Schwarzenegger) on the dimension of aggressiveness, but only when participants engaged in a self-reflective process (before the social comparison manipulation, participants were asked to reflect upon their own aggressiveness and consider how aggressive they are for about 1 min).

Results of these experiments were consistent with the modern view of social comparison as a process that can operate in the absence of awareness. In these experiments, all effects were predicted and all were in the expected direction, confirming the researchers' hypotheses. Although these results look quite convincing, it is now acknowledged that series of successful experiments with low power are very difficult to replicate in exact replications (Francis, 2012; Schimmack, 2012). Because it is impossible to know how many failed experiments, if any, researchers have conducted concerning this question, the true effect size is difficult to estimate, and thus, the probability to replicate these findings is unknown. Recently, researchers have developed indices of scientific quality to assess the extent to which a set of experiments has evidential value (the p -curve) and is replicable (the R -index) (Schimmack, 2014; Simonsohn, Nelson, & Simmons, 2014). We have used these indices to check the statistical results of Mussweiler et al. (2004). All analyses and results are available at <https://osf.io/gjc5b/>. The p -curve distribution is left-skewed, and the full p -curve test is significant, $Z = -1.69$, $p = 0.046$, indicating that this set of studies lacks evidential value. This suggests that direct replications of these studies are not expected to succeed. Consistent with this, the R -index test revealed that with a median observed power of 61%, and a success rate of 100%, the inflation rate was 38%. The R -index score was thus of 22%. According to the R -index manual, an R -index of 22% is consistent with a set of studies in which the null hypothesis is actually true. Thus, the p -curve and R -index indices were not encouraging. According to these indices, the set of experiments reported by Mussweiler et al. (2004) does not provide much evidence for the modern view of social comparison. However, it should be noted that these indices must be considered with caution, because they are new, because some of them are not published, and because they are controversial to some extent. Also, we underline that there is nothing specific to Mussweiler et al.'s studies in relation to this issue. In fact, most of previous experimental research on social comparison conducted before the crisis of confidence in social psychology would today be regarded as insufficiently sound from a methodological viewpoint.

In sum, research conducted so far on the automaticity of social comparison has produced mixed results. To date, it thus remains unclear whether or not social comparison is an efficient process, and whether or not it can operate without awareness.

1.3. The present research

In the two experiments reported here, we sought to further test the hypothesis that social comparison can operate without awareness (Mussweiler et al., 2004). We conducted these experiments in accordance with the new guidelines on good research practices (e.g., Funder et al., 2014). More precisely, we determined our sample size on the basis of a power analysis; we pre-registered our hypothesis; we focused on effect size rather than on p -value; we ran a close replication; we analyzed our data with both null hypothesis significance testing (NHST) and Bayesian analyses; and we posted our material and data on a public

repository (<https://osf.io/gjc5b/>).

In addition, our experiments were designed to address four main limitations of previous research. *First*, Mussweiler et al. (2004) did not explain how they determined their sample size. In contrast, we planned Experiment 1 to have 80% power to detect a significant effect (Cohen, 1992), based on an a priori power analysis (see the Method section). In Experiment 2, we increased statistical power to 95% to have a good chance (76%) to find significant effects in the two consecutive experiments (Schimmack, 2012). *Second*, Mussweiler et al. (2004) used a funneled debriefing method (Bargh, Chen, & Burrows, 1996) to test for participants' awareness of the primes. This method is based on verbalizations of subjective conscious experiences. To avoid the limitations of introspection and verbal reports, we used an objective awareness check, based on a measure of forced-choice discriminations. In that way, we could quantify the evidence in support of the null hypothesis that participants have not seen the subliminal primes using Bayesian analyses (Rouder, Speckman, Sun, Morey, & Iverson, 2009). *Third*, Mussweiler et al. (2004) conducted conceptual replications (using different manipulations and different dependent variables) rather than close replications (using the same manipulations and dependent variables). Conceptual replications are useful to test the same idea in different ways. However, close replications are especially important to verify the reliability of research findings (Simons, 2014). We designed Experiment 2 as a close replication of Experiment 1. It was not an exact replication because there were some procedural differences between the two experiments. In particular, we used a between-participants design in the first experiment, and a within-participant design in the second one. However, we used the exact same manipulation and the same dependent variable in the two experiments. *Fourth*, there was no standardized procedure in previous research that could be used by others to independently replicate the findings. Researchers have used many different procedures. This is problematic because it does not facilitate comparison between the findings of different experiments, or the replication of published findings. In contrast, we provide a replication package including the script of our experiment and all our stimuli (<https://osf.io/gjc5b/>). Anyone interested can download this script and use it to independently replicate our procedure.

In the present experiments, we used a classical manipulation of social comparison. Female participants were exposed to media images of the thin-ideal female body, which represents an extremely high standard of comparison for most women. Several meta-analyses indicate that women report greater body dissatisfaction, lower self-esteem, greater body appearance anxiety, and greater desire to be thin following acute exposure to media images of ultra-thin women, compared to media images of average size women or no-model images (Grabe, Ward, & Hyde, 2008; Groesz, Levine, & Murnen, 2002; Want, 2009). In the present research, we attempted to replicate some of these effects. In contrast to previous research, however, in our experiments we used a subliminal rather than a supraliminal presentation of the thin-ideal female body. We reasoned that if social comparison takes place without awareness, then subliminal exposure to media images of ultra-thin women would lead female participants to evaluate their own body more negatively. To the best of our knowledge, this hypothesis has not been tested before in high-powered experiments. In a previous experiment, Brown and Dittmar (2005) showed that brief exposures (10 s) to media images of the thin ideal are sufficient to increase body appearance anxiety in women, compared to a no model control condition. Here we examined whether we could find similar effects with masked images presented 500 times faster (20 ms).

In Experiment 1, we tested the hypothesis that female participants would report greater body appearance anxiety, drive for thinness, and body-specific self-discrepancies following subliminal exposures to media images of thin women, compared to media images of average size women or no images. In Experiment 2, we attempted to replicate Experiment 1 using a within-participant design rather than a between-participants design. In the two experiments, we included a number of

variables that have been shown in the past to moderate the effects of thin ideal exposure or the effects of subliminal primes. In Experiment 1, the moderating variables were thin-ideal internalization, body dissatisfaction, drive for thinness, social comparison tendency, and self-monitoring. In Experiment 2, the moderating variables were body dissatisfaction and self-monitoring.

2. Experiment 1

In this experiment, we tested the hypothesis that 20 ms exposures to media images of ultra-thin women are sufficient to increase body appearance anxiety, body-specific self-discrepancies, and drive for thinness in women. In a no-models control condition, the media images were simply omitted, preventing any social comparison. In another control condition, images of ultra-thin women were replaced by images of the same models, airbrushed to have a more realistic body shape, closer to average (a professional photographer retouched the images using Photoshop®). Because there were few differences between the ultra-thin and average size models conditions (see Appendix A for an illustration, and Supplementary Online Material for all images), we reasoned that the average size models could still represent a high standard of comparison for some women (e.g., body dissatisfied women). Thus, we expected a linear trend among the three conditions. Body appearance anxiety, body-specific self-discrepancies, and drive for thinness would be highest in the thin-ideal condition and lowest in the no-model control condition. The average size model condition was expected to fall in between. A significant linear trend would provide evidence that social comparison operates automatically and without awareness. Before looking at the data, we preregistered our hypotheses on the OSF (<https://osf.io/gjc5b/>).

Previous research has identified a number of individual differences variables that moderate effects of exposures to media images depicting the thin ideal: body dissatisfaction (Posavac, Posavac, & Posavac, 1998), self-monitoring (Henderson-King & Henderson-King, 1997), social comparison tendency (Dittmar & Howard, 2004), drive for thinness (Pavelo, 2006), and thin-ideal internalization (Brown & Dittmar, 2005). We expected women scoring high on these variables to be most affected by subliminal exposures to the thin ideal. We included these variables at the beginning of the procedure, before the subliminal social comparison manipulation. This was done to test for possible moderation effects, but also to ensure that participants engaged in self-reflection on their own body appearance before the social comparison manipulation. Previous research suggests that self-reflection might be necessary to observe self-evaluative consequences of subliminal social comparison (Mussweiler et al., 2004, Study 3).

2.1. Method

We report all measures, manipulations, and exclusions in these experiments in text or footnotes.

2.1.1. A priori power analysis

We conducted an a priori power analysis to determine the required sample size for this experiment. We designed this experiment to have 80% power to detect a small-to-medium effect size (Cohen's $d = 0.42$), corresponding to the mean effect size in social psychology¹ (cf. Richard

¹ As suggested by Dr. Want, who worked as a reviewer on this paper, we may have chosen a different effect size. For example, we may have based the expected effect size on previous meta-analytic estimates of supraliminal exposure to thin-ideal images (i.e., Cohen's $d = 0.35$, cf. Want, 2009) rather than on the mean effect size in social psychology. We used the mean effect size in social psychology as a rule of thumb. As noted in the pre-registered protocol, it was used to provide 80% power to detect an effect as large as the one found by Mussweiler et al. (2004). Moreover, it was a conservative estimate of the expected effect size compared to relevant research by Mussweiler et al. (2004) and Brown and Dittmar (2005).

et al., 2003). In the most relevant experiments conducted so far, the observed effect size was larger than that. Brown and Dittmar (2005) found an effect size of 0.70 (Cohen's d) on body appearance anxiety following brief exposure (10s) to thin-ideal images. Mussweiler et al. (2004, Study 2) found an effect size of $d = 0.50$ of subliminal exposure with extremely high or low standards of comparison. Thus, the use of the mean effect size in social psychology ($d = 0.42$) as a prior was arguably a relatively conservative estimate of the expected effect size.

Because we had a specific (unidirectional) hypothesis, we set the Type I error rate at $p < 0.05$ (one-tailed). In that way, we had at least 80% power to detect marginal effects ($p < 0.10$) in two-tailed tests. This seemed justified because when in the expected direction, mean differences with marginal p values are usually interpreted as providing support for a specific hypothesis. We computed the required sample size using G*Power 3.1*. The required total sample size was estimated at 142 participants, so about 50 participants in each condition. Thus, we decided in advance to include a total sample size of 150 participants. Following Simonsohn's (2015) recommendations, this sample size is > 2.5 larger than the one used in previous experiments (Mussweiler et al., 2004). Therefore, it should provide 80% power to detect an effect the previous experiments had only 33% power to detect.

2.1.2. Participants

A total of 150 French women took part in this experiment (mean age = 27.25 years, $SD = 11.33$ years). Three female research assistants recruited them among the students of the University of Poitiers, France. Upon arrival to the laboratory, participants were welcomed and escorted into one of three identical experimental boxes (for a video of the lab setup, see <https://osf.io/4vy3w/>). Due to a technical problem (dropped frames) with one of the computers used for data collection, we had to exclude data for 23 participants, leaving 127 participants in the final sample. The reduction in sample size caused a decrease of power (76%). However, statistical power in this experiment was still much larger than in previous research and provided reasonable chances of detecting significant effects.

2.1.3. Procedure

Participants were seated in front of a PC computer running on Windows with a CRT monitor (60 Hz, frame rate = 16.67 ms). We used a standardized computer-administered procedure. After completion of an informed consent form, participants were asked to read and follow the instructions presented on the computer screen. The script for this experiment was built using PsychoPy* (Peirce, 2007, 2009) and is available on the OSF (<https://osf.io/gjc5b/>). All the instructions were given to participants via a computer. Participants were informed that they were taking part in an experiment on body image. They first completed self-report measures (moderators). Next, they took part in an ostensible personality test (used for subliminal priming). Then, they completed the dependent variables, and an objective awareness check measure.

2.1.4. Moderators

For the assessment of each moderator, a validated scale was used. Participants indicated their responses to each item of each scale on a 5-point Likert scale (1 = totally disagree; 5 = totally agree). The different scales used here² have been shown to have good psychometric properties in prior research. For each scale, responses were averaged across items to form a single score indicating the corresponding construct. Thus, average scores on each scale could range from 1 to 5.

Individual differences in thin-ideal internalization were assessed using the 8-item internalization subscale of the sociocultural attitudes

² The script also included 3 items used to assess consumerism (e.g. "I'd like to buy clothes that make me look thinner"). They were included for explorative purposes. Thus, they are not discussed further in this paper.

towards appearance questionnaire (Heinberg, Thompson, & Stormer, 1995; French version by Rousseau, Valls, & Chabrol, 2010). Representative items were: “I believe that clothes look better on thin models” and “Photographs of thin women make me wish I were thin.” Body dissatisfaction was assessed with the scale developed by Garner, Olmstead, and Polivy (1983) (French version by Chatard & Selimbegović, 2011). This scale included 9 items (e.g., “I think that my thighs are too large”; “I like the shape of my buttocks”). Drive for thinness was assessed using 6 items (e.g., “I am preoccupied with the desire to be thinner”; “I think about dieting”) taken from Garner et al. (1983) (French version by Rousseau, Knotter, Barbe, Raich, & Chabrol, 2005). The original scale included 7 items, but we did not use the first item, which had very low inter-item correlation in validation studies (Garner et al., 1983). Body comparison tendencies were assessed using the 5 items (e.g., “At parties or other social events, I compare my physical appearance to the physical appearance of others”; “In social situations, I sometimes compare my figure to the figures of other people”) of the physical appearance comparison scale (Thompson, Heinberg, & Tantleff-Dunn, 1991; French version by Dany & Urdapilleta, 2012). Finally, participants completed Snyder's (1974; French version by Gana & Brechenmacher, 2001) self-monitoring scale (25 items, e.g., “I have trouble changing my behavior to suit different people and different situation”; “My behavior is usually an expression of my true inner feelings, attitudes, and belief”), assessing the degree to which people adjust their behavior to situational demand.

2.1.5. Subliminal social comparison

Next, participants completed a subliminal affective priming (SAP) task (Murphy & Zajonc, 1993). In this task, participants were asked to evaluate 28 Chinese ideograms, presented in a random order, by pressing the “I” key for “I like it”, and the “E” key for “I dislike it” on an AZERTY keyboard. The 28 ideograms were presented twice each, for a total of 56 trials. This task was used to induce subliminal social comparisons. There were 3 different versions of this task, which differed only by the subliminal image that preceded each ideogram (see Fig. 1). In the thin ideal models condition, each ideogram was preceded by a media image of an ultra-thin woman, presented for 20 ms. The thin ideal image was “sandwich masked” (using pre- and post masks) with a heavily scrambled version of the intact image (see Fig. 1). The pre- and post-masks were also presented for 20 ms. We used 14 different media images of the thin ideal body (see Supplementary Online Material), each presented four times. In the average size models condition, each ultra-thin model image was replaced with an equivalent average size model image. A professional photographer retouched the 14 ultra-thin models images to subtly change their physical appearance and make them look more realistic (or average size). Finally, in the no-model image condition, the ideograms were not preceded by any body image. Participants were randomly assigned to one of the three conditions.³

It is noteworthy that the masks used in this experiment (see Fig. 1) were created in a way to optimize the appropriate masking of body shape and edges. Because we were mainly interested in the visual processing of body shape, involving some specific brain areas (Kim, Biederman, Lescroart, & Hayworth, 2009), it was important to contrast stimuli with a defined shape (the thin ideal vs. average size models) with those that do not have one (heavily scrambled masks) (see Stojanoski & Cusack, 2014). Thus, we used a box scrambling masking method (Vogels, 1999) to create the masks. Each intact image was divided into an equal number of independent boxes (at the pixel level),

and each box was randomly repositioned to a new location in the image. That box-scrambling method is recommended in experiments that focus on specific object properties, such as shape and edges (Stojanoski & Cusack, 2014).

Before data collection, we used the black box tool kit* (Plant, 2016) to certify the very short duration of our subliminal primes, independently from the PsychoPy* software. The frame rate was 16.67 ms ($SD = 0.08$) on average. Thus, all primed images were presented for < 20 ms. During data collection, we recorded all frames such that we could check that we did not drop frames during the presentation of stimuli. One computer, which worked perfectly fine at the beginning of data collection, started to drop frames at some point (some of the mask and primes were not presented). Therefore, we excluded the observations (15.33%) for which a consistent frame rate could not be recorded. In a recent series of seven experiments, we have used the SAP task to subliminally prime participants with different images (Chatard, Hirschberger, Faure, Champeaux, & Mermillod, 2017). An internal meta-analysis of the 7 experiments clearly showed that this task could be successfully used to prime subliminal images.

2.1.6. Dependent variables

After the SAP task, participants completed 3 measures used to assess participants' self-evaluations following social comparisons with the thin ideal. First, they completed a measure of discrepancy between actual and ideal body images, adapted from the body size guide (Harris, Bradlyn, Coffman, Gunel, & Cottrell, 2008). They were presented with pictograms of a female body with different shapes and asked to select which one corresponded the most to their actual body shape. Participants pressed a response key from “A” to “J” to indicate their responses. Next, the pictograms were presented a second time, and participants were asked to select which one corresponded the most to their ideal body shape. Responses were recoded from $J = 1$ to $A = 10$, with higher scores indicating thinner self-perceptions. An actual-ideal body discrepancy index was computed by subtracting the actual body perception from the ideal body perception. Scores on the discrepancy index could range from -9 to 9 . Positive scores indicated that participants wanted to be thinner than they are, while negative scores indicated that participants wanted to be larger than they are.

Second, participants completed a state measure of body appearance anxiety (taken from Reed, Thompson, Brannick, & Sacco, 1991, and also used by Brown & Dittmar, 2005), assessing how anxious and nervous participants feel “right now” about their weight and various body parts. This scale included 8 items: “Right now, I am thinking that I am too fat and it makes me feel anxious”; “Right now, I find my buttocks too large and it makes me feel nervous”; “Right now, I find my hips too large and it makes me feel nervous”; “Right now, I find my thighs too large and it makes me feel nervous”; “Right now, I think my stomach is too big and it makes me feel nervous”; “Right now, I think my arms are too big and it makes me feel anxious”; and “Right now, I am thinking about my weight and it makes me feel anxious”.

Third, participants completed a state measure of the desire for thinness (6 items: “Right now, I would like to be thinner”; “Right now, I would like to lose weight”; “Right now, I would like to diet”; “Right now, I would like to have clothes that make me look thinner”; “Right now, I would like to buy body care products that make me look thinner”; and “Right now, I would like to have a cosmetic surgery to look thinner”). This measure was built for the purpose of this experiment. It was adapted from existing scales to assess momentary, acute desire for thinness, rather than a trait measure of this construct.

2.1.7. Objective awareness check

Finally, participants were presented again with 10 Chinese ideograms. This time, they were informed that a subliminal image was presented before some of the ideograms. On each trial, participants were asked to guess whether or not a subliminal image was presented. On 5 trials, there was a masked subliminal image of an ultra-thin

³ The script also included an implicit measure of body image. Following the subliminal social comparison manipulation, participants completed another block of the SAP task (28 trials). This time the body images were replaced by subliminal words (e.g., “me”, “my body”, “my thighs”), masked by a rapid succession of black and white screens. This task was included for explorative purposes. It has not been validated before. Thus, it is not discussed further in this paper.

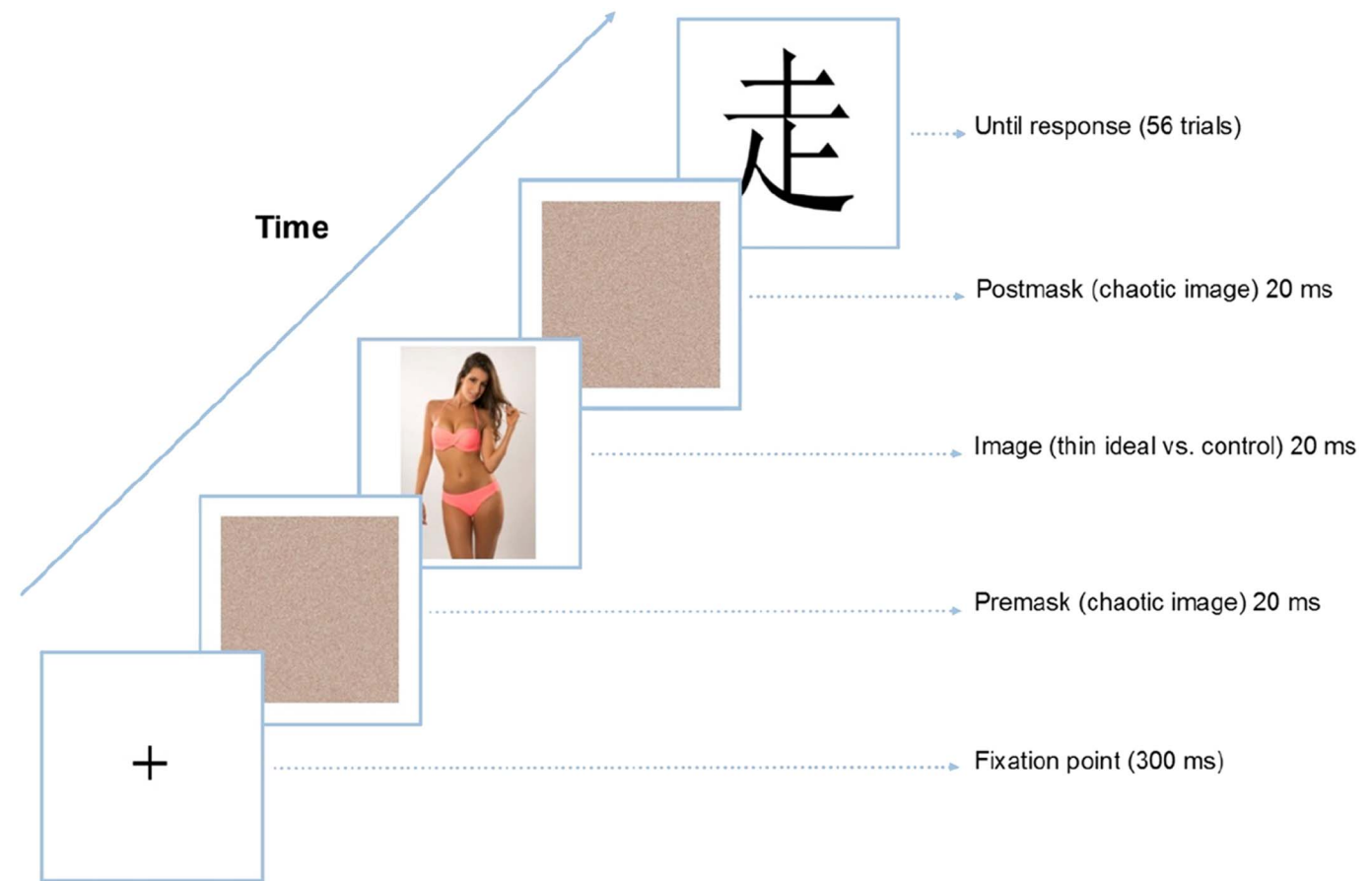


Fig. 1. Illustration of the subliminal procedure used to prime body images in Experiments 1 and 2.

woman (Fig. 1). On the 5 remaining trials, there was no image. The trials were presented in a fully random order and participants pressed the “I” key if they thought there was a subliminal image and the “E” key if they thought there was no subliminal image.

2.2. Results

2.2.1. Preliminary analyses

The descriptive statistics and correlations are reported in Table 1. In preliminary analyses, we checked the randomization of conditions by examining possible effects of conditions on the moderating variables. In analyses of variance, there were no effects of conditions on these variables ($ps > 0.18$), confirming the appropriate randomization of conditions. A detailed description of all analyses and effects is reported in Supplementary Online Material.

Table 1
Descriptive statistics.

	Alpha	Mean	SD	1	2	3	4	5	6	7	8
1. Thin internalization	0.94	2.57	1.06	–	0.25**	0.47***	0.62***	0.28**	0.12	0.37***	0.39***
2. Body dissatisfaction	0.88	3.38	0.99		–	0.62***	0.28**	–0.08	0.66***	0.81***	0.62***
3. Drive for thinness	0.90	3.00	1.19			–	0.52***	0.16	0.48***	0.77***	0.78***
4. Social comparison tendencies	0.88	3.41	1.09				–	0.39***	0.19*	0.40**	0.50***
5. Self-monitoring	0.76	2.67	0.48					–	–0.21*	0.04	0.10
6. Discrepancy index	–	1.57	1.64						–	0.65***	0.61***
7. Body appearance anxiety	0.94	2.89	1.26							–	0.81***
8. Desire for thinness	0.89	2.82	1.20								–

Note: N = 127.
*** $p < 0.001$.
** $p < 0.01$.
* $p < 0.05$.

2.2.2. Awareness check

Participants reported that a subliminal image was presented before the ideograms on 8.34% ($SD = 22.73$) of the trials when there was a subliminal image, and on 8.66% of trials ($SD = 25.01$) when there was no subliminal image. In a paired-sample t -test, the hit and false alarm rates were not different, $t(126) = -0.40$, $p = 0.68$, $d = -0.03$. We used a Bayesian t -test, with the default Cauchy prior width of 0.707 (Wetzels, Raaijmakers, Jakab, & Wagenmakers, 2009), to test the hypothesis that the hit rate was greater than the false alarm rate. In this analysis, we found strong support for the null hypothesis, $BF_{01} = 13.57$. Given the data, the null hypothesis was 13.57 times more likely than the alternative hypothesis. Although Bayes factors provide a continuous measure of evidence, Bayesian statisticians often used criterion values to facilitate decision-making. Following Jeffreys (1961), Bayes factors (BF_{01}) > 3 (or 10, or 30, respectively) are usually

considered as providing substantial (or strong, or very strong, respectively) evidence for the null over the alternative hypothesis. Thus, in this experiment we found strong support for the hypothesis that participants could not detect the presence of subliminal images above chance level.⁴

2.2.3. Main effects

We closely followed the pre-registered protocol (<https://osf.io/uz9si/>) to analyze the data. We first examined effects of the conditions on our dependent variables. The means and standard deviations are presented in Table 2. As shown in this table, the means were generally consistent with our predictions: actual-ideal discrepancies, body appearance anxiety, and desire for thinness were higher in the ultra-thin models condition than in the no-models condition. However, the means in the average size models condition were not intermediate, but very similar to the means in the no-models control condition. We used two orthogonal contrasts to test our pre-registered hypotheses. The first one (linear contrast) contrasted the ultra-thin condition (coded +1) and the no-models condition (coded -1). The average size models condition was coded 0. The second contrast tested the residual difference (the average size models condition was coded +2, and the two other conditions were coded -1). All reported *p*-values are two-tailed, but we interpret marginal effects as significant when clearly predicted in advance in the pre-registration. A detailed description of all analyses and effects is reported in Supplementary Online Material.

On the *discrepancy* index, we found no effect of the linear contrast, $\beta = 0.14$, $t(123) = 1.60$, $p = 0.11$, and no effect of the residual contrast, $\beta = -0.12$, $t(123) = -1.38$, $p = 0.17$. Contrary to our expectations, participants in the ultra-thin condition, compared to those in the no-models condition, did not evaluate their actual body shape as being larger than what they would like in the ideal. Results of further (not pre-registered) analyses revealed that the linear contrast approached significance on the rating of the actual body image, $\beta = -0.16$, $t(123) = -1.86$, $p = 0.065$, but not on the rating of the ideal body image, $\beta = -0.11$, $t(123) = -1.27$, $p = 0.20$.

On *body appearance anxiety*, we found a significant effect of the linear contrast, $\beta = 0.19$, $t(124) = 2.15$, $p = 0.033$, and no effect of the residual contrast, $\beta = -0.07$, $t(124) = -0.89$, $p = 0.37$. Consistent with our (pre-registered) hypothesis, participants in the ultra-thin condition, compared to those in the no-models condition, reported greater body appearance anxiety.

On the *desire for thinness*, we found no significant effects of the linear contrast, $\beta = 0.06$, $t(124) = 0.76$, $p = 0.44$, or of the residual contrast, $\beta = -0.06$, $t(124) = -0.69$, $p = 0.49$.

2.2.4. Moderation effects

Following the pre-registered protocol, we tested whether the effects of conditions on our dependent variables were moderated by individual differences in thin ideal internalization, body dissatisfaction, drive for thinness, social comparison tendency, and self-monitoring. Each dependent variable was regressed on the linear contrast, the residual contrast, each moderator variable (mean centered), and the product terms between the linear contrast and each moderator, and between the residual contrast and each moderator. Given the large number of effects tested, we report in what follows a summary of the main findings, focusing on the predicted moderation effects.

On the measure of *discrepancy* between actual and ideal body

images, none of the predicted interactions were significant. The effect of condition (linear contrast) was not moderated by individual differences in thin-ideal internalization, body dissatisfaction, drive for thinness, social comparison tendency, and self-monitoring (see Supplementary Online Material). The main effects of body dissatisfaction ($p < 0.001$) and drive for thinness ($p = 0.019$) were significant. There was also a non-predicted interaction between the residual contrast and body dissatisfaction ($p = 0.03$).

On the measure of *body appearance anxiety*, none of the predicted interactions were significant. The effect of conditions (linear contrast) was not moderated by individual differences in thin-ideal internalization, body dissatisfaction, drive for thinness, social comparison tendency, and self-monitoring (see Supplementary Online Material). The effect of condition (linear contrast) was significant ($p = 0.019$). The main effects of body dissatisfaction ($p < 0.001$) and drive for thinness ($p < 0.001$) were also significant.

On the measure of *desire for thinness*, none of the predicted interactions were significant. The effect of condition (linear contrast) was not moderated by individual differences in thin-ideal internalization, body dissatisfaction, drive for thinness, social comparison tendency, and self-monitoring (see Supplementary Online Material). The main effects of body dissatisfaction ($p < 0.03$), drive for thinness ($p = 0.001$), and social comparison tendency ($p < 0.02$) were significant.

2.2.5. Supplementary (not preregistered) analyses

We conducted Bayesian independent sample *t*-tests, with the default Cauchy prior width of 0.707, to quantify the evidence in favor of our specific hypotheses. Because we predicted a linear trend between the three conditions, we excluded the average size models condition from these analyses, and directly tested the mean difference between the ultra-thin models condition and the no-models condition. We expected body discrepancy, body appearance anxiety, and desire for thinness to be higher in the ultra-thin models condition, compared to the no-models condition. We found substantial support for this hypothesis on body appearance anxiety, $BF_{10} = 3.45$, but not on body discrepancy, $BF_{10} = 1.13$, or desire for thinness, $BF_{10} = 0.45$.

2.3. Discussion

In this experiment, we found that women reported greater body appearance anxiety following subliminal exposures to the thin ideal body, compared to a no-models control condition (Cohen's $d = 0.47$). Using an objective awareness check and Bayesian analyses, we found strong evidence that participants were unaware of the subliminal primes. Yet, they reported being significantly more anxious and nervous about their body weight following subliminal exposures to the thin ideal models. Clearly, this suggests that participants engaged in implicit social comparison with the ultra-thin models, and felt bad about their own body image as a consequence. From a theoretical perspective, the present findings entail that social comparison takes place without awareness and affects explicit self-evaluation.

The present findings are consistent with previous experiments using supraliminal presentation of thin ideal body images (Brown & Dittmar, 2005). They are also consistent with past research showing contrast effects in self-evaluation following subliminal exposures to extremely high standards of comparison (Mussweiler et al., 2004). However, because we pre-registered our hypotheses and used a much larger sample than in any previous experiments on subliminal social comparison, the present findings provide the strongest available evidence in favor of the automaticity of social comparison, due to its unconscious nature. The effect size was rather small, but it fell in the range of the typical effects studied in social psychology (i.e., small-to-medium effect sizes). Using Bayes factors as measures of evidence, we found substantial evidence for our hypothesis on body appearance anxiety. Thus, we believe that the present findings are credible and that they provide

⁴ Because the choice of a prior influences Bayes-Factors, we decided in advance to use a default prior in this research. However, we agree that the default Cauchy prior width of 0.707 is not the most appropriate prior for the awareness check, as it assumes a median population effect size of $d = 0.70$ under the alternative hypothesis. Yet, it is noteworthy that the results were very similar when we set a smaller median effect size under the alternative hypothesis. With a Cauchy prior width of 0.20, corresponding to a small median effect size, we still found substantial evidence for the null hypothesis (i.e., participants could not see the subliminal primes), $BF_{01} = 4.23$.

Table 2
Means, standard errors, and effect sizes.

Condition	No-models (n = 50)	Average size models (n = 39)	Ultra-thin models (n = 37)	Mean difference (ultra-thin – no-models)	95% confidence interval	Cohen's <i>d</i>
Discrepancy index	1.46 (0.23)	1.30 (0.26)	2.02 (0.26)	0.56 (0.39)	[– 0.17; 1.31]	0.32
Body appearance anxiety	2.70 (0.17)	2.77 (0.19)	3.28 (0.20)	0.57 (0.26)	[0.05; 1.10]	0.47
Desire for thinness	2.78 (0.17)	2.71 (0.19)	2.97 (0.19)	0.19 (0.25)	[– 0.30; 0.69]	0.16

new evidence for the automaticity of social comparison (Gilbert et al., 1995).

At the same time, there were some surprising and unexpected effects, and most of our hypotheses were not confirmed by the data. First, although the means in the average size models conditions did not deviate from a linear trend, they were very similar to the means in the no-models control condition, and different from the means in the ultra-thin models condition. We did not expect that because there were very few differences between the average size models and the ultra-thin models (see Supplementary Online Material). However, this suggests that participants could discriminate subtle differences between the two kinds of models, even at the implicit or unconscious level. Second, we found no significant effects of subliminal social comparison on body discrepancies and desire for thinness, although the means were in the expected direction. Third, we found no evidence that effects of subliminal social comparisons were moderated by individual difference variables. Several potential moderators were considered in this experiment. Yet, none was significant to predict our dependent variables.

In sum, the results of Experiment 1 provided support for the main hypothesis of this paper that subliminal social comparison affects explicit self-evaluation. However, the effects were significant only on one out of the three dependent variables, and we found no support for moderation. Given the large number of effects tested here, it is still possible that the significant effect observed on body appearance anxiety was obtained by chance alone. Thus, in Experiment 2, we adopted a purely confirmatory approach and attempted to replicate the main effect of conditions found here on body appearance anxiety specifically. If Experiment 1 provided a reliable and a reasonably precise estimate of the effect size of subliminal social comparisons on body appearance anxiety, then we should be able to replicate this effect in a close replication.

3. Experiment 2

Our aim in this second experiment was twofold. First, we attempted to provide further evidence either in favor or against the hypothesis that social comparison takes place without awareness and thus automatically affects explicit self-evaluations (Mussweiler et al., 2004). Second, ideally we wanted to find a robust effect of subliminal social comparison on self-evaluation, which could be easily replicated in independent replications. In an effort to reach these goals, we decided to shift from a between-participants design to a within-participant design. A within-participant design allows testing our hypothesis in a high-powered experiment but with a much smaller sample, thereby facilitating independent replications. The use of a within-participant design might also provide converging evidence for our hypothesis, showing within-individual changes rather than between-individuals differences in self-evaluation following subliminal social comparison with a high standard. In Experiment 2, we simplified the design while trying to keep all essential parts of the procedure constant. Thus, we included only two moderating variables: body dissatisfaction and self-monitoring. Using a purely confirmatory approach, we also included only one dependent variable (body appearance anxiety), which was completed twice by the same participants.

We did not pre-register this experiment but our hypotheses were the same as in the previous experiment. Female participants should report

greater body appearance anxiety when subliminally primed with ultra-thin models (in the ultra-thin model condition), than when they are not primed with such extreme standards of thinness (in the no-model condition).

3.1. Method

We report all measures, manipulations, and exclusions in this experiment in text or footnotes.

3.1.1. A priori power analysis

Using G*Power 3.1*, we estimated the required sample size to replicate the effect of conditions (no-models vs. ultra-thin models) found on body appearance anxiety in Experiment 1 ($d = 0.47$, see Table 2). We set the Type I error rate (α) at 5% (one-tailed), and power ($1 - \beta$) at 95%. The required total sample size to replicate an effect size of 0.47, in a paired sample *t*-test, was estimated at 51 participants. Thus, we decided in advance to include 51 participants in this experiment.

3.1.2. Participants

Participants were 51 undergraduate female students recruited from our research pool ($M_{\text{age}} = 20.49$ years, $SD = 3.25$ years). They signed an informed consent form before the experiment and took part in exchange for credit in their introductory psychology course.

3.1.3. Procedure

Participants were seated in front of a PC computer running on Windows with a CRT monitor (85 Hz, frame rate = 11.76 ms). The procedure was very similar to the one used in Experiment 1, except that we used a within-participant design rather than a between-participants design. The script of the experiment is available from the OSF (<https://osf.io/gjc5b/>). Participants were informed that they were to take part in an experiment on body image. They first completed the body dissatisfaction and the self-monitoring scales (response scales ranged from 1 = “do not at all agree” to 7 = “completely agree”, responses were averaged). Next, they completed the SAP task, and the body appearance anxiety measure (response scales ranged from 1 = “do not at all agree” to 5 = “completely agree”, responses were averaged). After that, they completed the SAP task and the body appearance anxiety measure for a second time. As in Experiment 1, the SAP task was used to prime participants with subliminal images of ultra-thin women (see Supplementary Online Material). The subliminal images were presented either on the first or the second SAP task, based on a computer randomized order. When the subliminal images were not presented, participants were not exposed to any models. Finally, participants completed an objective awareness check, similar to the one used in Experiment 1, but with 20 trials. At the end of the experiment, which took about 10 min to complete, participants were thanked and fully debriefed.

3.2. Results

In what follows, all reported *p*-values are two-tailed. However, because our hypothesis was clearly unidirectional, we interpret marginal effects (if any) on our main dependent variable as significant when in

Table 3
Descriptive statistics.

	Alpha	Mean	SD	1	2	3	4
1. Body dissatisfaction	0.85	3.87	1.30	–	– 0.09	0.86***	85***
2. Self-monitoring	0.70	3.92	0.53	–	–	– 0.03	0.00
3. Body appearance anxiety (no models)	0.93	2.26	1.14	–	–	–	0.97***
4. Body appearance anxiety (ultra-thin models)	0.93	2.37	1.15	–	–	–	–

Note: N = 51.

*** $p < 0.001$.

the predicted direction.

3.2.1. Preliminary analyses

The descriptive statistics and correlations are reported in Table 3. In preliminary analyses, we checked for possible contamination effects by testing whether the order of the two SAP tasks influenced the responses on the body appearance measures. In a 2 (order of presentation of the subliminal images: first time vs. second time) \times 2 (body appearance condition: no-models vs. ultra-thin models) mixed ANOVA, with repeated measures on the last factor, we found no significant interaction between the two factors, $F(1, 49) = 0.008$, $p = 0.93$, $\eta^2 < 0.001$. Thus, we found no evidence that responses on the body appearance anxiety measures (in the no-models and ultra-thin models conditions) were influenced by the order of the SAP task (whether the subliminal images were presented the first or the second time on the SAP task).

3.2.2. Awareness check

Participants reported that a subliminal image was presented before the ideograms on 21.37% ($SD = 23.71$) of trials when there was a subliminal image, and on 22.06% of trials ($SD = 23.43$) when there was no subliminal image. A paired sample t -test indicated that the hit and false alarm rates were not statistically different, $t(50) = -0.43$, $p = 0.66$, $d = -0.06$. We used a Bayesian paired sample t -test, with the default Cauchy prior width of 0.707, to test the hypothesis that the hit rate was greater than the false alarm rate. In this analysis, we found stronger evidence for the null hypothesis, $BF_{01} = 8.92$, than for the alternative hypothesis, $BF_{10} = 0.11$. Indeed, the null hypothesis was about 9 times more likely than the alternative hypothesis, providing substantial support for the null hypothesis.⁵

3.2.3. Main analyses

As shown in Table 3, the mean of body appearance anxiety was slightly higher in the ultra-thin models condition than in the no-models condition (mean difference = 0.11, $SE = 0.03$, $CI_{95} [0.044; 0.176]$). A paired sample t -test revealed that this difference was significant, $t(50) = 3.36$, $p = 0.001$, $d = 0.47$, with the same effect size as the one found in Experiment 1. In a Bayesian paired sample t -test, with the default Cauchy prior width of 0.707, the present data provided very strong evidence for the hypothesis that subliminal exposures to the thin ideal increased body appearance anxiety, compared to the no-models condition, $BF_{10} = 39.93$.

3.2.4. Complementary analyses

In Experiment 2, we found no evidence that individual differences in body dissatisfaction moderated the effect of conditions on body appearance anxiety (for the details of the results, see Supplementary Online Material). As shown in Table 3, the correlations between body dissatisfaction and body appearance anxiety were very strong and not different in the two exposure conditions. In the same way, we found no

evidence that self-monitoring moderated the effects of conditions on body appearance anxiety. The effects of condition on body appearance anxiety remained significant after control for body dissatisfaction and self-monitoring.

3.3. Discussion

In this close replication of Experiment 1, we successfully replicated our previous findings using a within-participants rather than a between-participants design. Subliminal exposure to the thin ideal caused transient changes in body appearance anxiety within the same individuals. This finding confirmed that the effects found in Experiment 1 were not spurious. The probability that the same effect was found by chance in two consecutive experiments is quite low (i.e., the product of the two p -values, so about one chance in 30,303). Consistent with this, using Bayes factors as measures of evidence, we found very strong evidence in favor of the hypothesis. Indeed, the present data were about 40 times more likely under our hypothesis than under the null hypothesis. These findings are consistent with the modern view of social comparison as an unconscious, and thereby an automatic process. Because we also found substantial evidence that participants were unaware of the subliminal primes, the present findings support the conclusion that social comparison takes place without conscious awareness and affects explicit self-evaluation.

4. General discussion

The idea that social comparison is automatic is not only central to contemporary social comparison research, but it is also quite popular in social psychology. For example, the paper by Mussweiler et al. (2004) on subliminal social comparison is one of the most downloaded articles from the *Journal of Experimental Social Psychology* website.⁶ As we have seen, however, this article, as well as some other relevant publications, does not provide solid evidence for the automaticity of social comparison. Indeed, this literature is plagued by retracted publications and low-powered experiments. At a time when most researchers doubt the reliability of social psychological findings, it is especially concerning that one of the central assumptions of modern social comparison research is not supported by more solid evidence. Adopting a skeptical approach, we tested whether social comparison can take place outside of conscious awareness. Across two experiments, we found converging evidence that subliminal social comparison affects how individuals consciously evaluate themselves. Women exposed to non-visible media images of ultra-thin women reported being more nervous and anxious about their own weight. In what follows, we discuss why we believe it is an important finding; we acknowledge some limitations of the present experiments and discuss new avenues of research; and, we conclude by making recommendations for future research on social comparison.

4.1. Theoretical implications

The present findings contribute to research and theorizing on social comparison in an important way. Currently, most social psychologists appear to assume, following Gilbert et al. (1995), that social comparison is automatic (Want et al., 2015). Nonetheless, the empirical evidence supporting this claim is quite limited. Moreover, recent research has challenged this claim, failing to find evidence that high cognitive load facilitates social comparison effects on explicit self-evaluation (Want et al., 2015; Want & Saiphoo, 2017). In contrast, the present findings corroborate the assumption that social comparison can operate without awareness, and that is therefore automatic according to this criterion (which also largely overlaps with the criteria of intention

⁵ With a Cauchy prior width of 0.20, corresponding to a small median effect size, we still found substantial evidence for the null hypothesis (i.e., participants could not see the subliminal primes), $BF_{01} = 3.03$.

⁶ Retrieved on December 6, 2016 from: <http://www.journals.elsevier.com/journal-of-experimental-social-psychology/most-downloaded-articles>.

and control in this case). They thus considerably increase our confidence in the modern view of social comparison as an automatic process (Mussweiler et al., 2004). In two consecutive experiments, using between- and within-individual designs, high-powered samples, pre-registered hypotheses, objective awareness checks, null-hypothesis significance tests and Bayesian analyses, we found substantial (Experiment 1) and very strong (Experiment 2) evidence that social comparison occurs without conscious awareness. Our findings thus strengthen the confidence in the reliability of previously published findings (Mussweiler et al., 2004). Subliminal exposure to extremely high standards of comparison lead to contrast effects; individuals evaluated themselves more negatively following such exposure (compared to different control conditions), even if they were unaware that this negative evaluation is triggered by social comparison processes. Our findings are at odds with those reported by Want et al. (2015) and Want and Saiphoo (2017), but they are consistent with recent research showing that high cognitive load facilitates social comparison processes when self-evaluation is assessed via implicit (automatic) measures (Bocage-Barthélémy et al., 2017). Social comparison appears to be an automatic process in the sense that it requires little cognitive effort (Gilbert et al., 1995), and little attention or awareness (Mussweiler et al., 2004).

The present findings are also important because they offer new insights on the effects of thin ideal exposure in social psychology. In many experiments, exposure to the thin beauty ideal has been found to increase body image concerns in women (e.g., Grabe et al., 2008). Social comparison theory offers a parsimonious account for this effect. Yet, because most experiments used relatively long durations of stimuli exposures, one possibility is that some of the evidence can be explained away as a result of demand characteristics in the experiment; when participants are exposed to media images of ultra-thin women, they could slightly change their usual behavior because they correctly guess the experimenters' hypothesis. As Jansen and de Vries (2002) wrote: "data from supraliminal studies may be seriously confounded by demand characteristics like answering in a socially desirable way" (p. 208). The problem of demand characteristics, which involves research participants being aware of what the researcher is investigating, is acknowledged in many articles on thin ideal exposures. However, the problem of demand characteristics has not been taken seriously until now. The results of the only study in which demand characteristics were manipulated by instructions (Mills, Polivy, Herman, & Tiggesmann, 2002) are particularly worrisome in this regard. They suggest that women are more likely to report that looking at thin media images makes them feel bad when demand characteristics are high rather than low. Thus, effects of thin ideal exposure may be greatly exaggerated in the literature due to demand characteristics. Because in the present experiments participants were unaware of being exposed to ultra-thin models, the findings are somewhat reassuring as they suggest that demand characteristics alone cannot explain the negative consequences of thin ideal exposures on body appearance anxiety. To the best of our knowledge, the present experiments are the first to clearly show this.

Finally, the present experiments contribute to recent debates on the replication crisis by showing that some previously published findings in social psychology can be replicated, at least conceptually, even if they have low indices of scientific quality (*p*-curve and *R*-index). In other words, low indices of scientific quality do not imply that the related findings are not valid. These indices are useful tools to estimate the strength of the evidence of a set of studies. However, the best way to check whether an effect exists still remains to test it experimentally using high-powered designs. The present experiments provide converging evidence that social comparison happens subliminally, and thus, automatically.

4.2. Limitations and future research

The present experiments have certain limitations that deserve

further attention in future experiments. First, further research is needed to delineate the boundary conditions of subliminal social comparison effects. As we have seen, our results are generally consistent with previous research by Mussweiler et al. (2004). In their last experiment, these researchers found that subliminal social comparisons affected self-evaluation only when participants were led to engage in a self-reflective process before the comparison. In our experiments, all participants completed (at least) the body dissatisfaction scale before the subliminal social comparison. Thus, in a sense, they all engaged in a self-reflective process before the comparison, and it is quite possible that our effects are restricted to this specific situation. At the same time, Mussweiler et al. (2004) did not induce a self-reflection process in their initial experiments but still found significant effects of subliminal social comparison on self-evaluation. Thus, it remains unclear whether or not a self-reflective process is a necessary condition to observe the effects we report here. Further study is needed to examine this issue.

Second, in the present experiments we predicted and found contrast effects following subliminal priming of thin ideal body images. Our findings are consistent with previous research on supraliminal exposures to thin ideal body images (Grabe et al., 2008), and with research on subliminal social comparison with extremely high standards of comparison (Mussweiler et al., 2004). However, the literature on prime-to-behaviors effects indicates that subliminal primes can also lead to assimilation effects (e.g., Bargh et al., 1996; Devine, 1989; Kawakami, Dovidio, & Dijksterhuis, 2003; Wheeler & Petty, 2001). For example, Bargh et al. (1996) found that participants who were subliminally primed with pictures of African American faces displayed more aggressive facial expressions following a frustrating event than did control participants. To date, there is no consensus in the literature on when and why subliminal primes lead to assimilation or contrast effects. Mussweiler's (2003) selective accessibility model suggests that subliminal primes are assimilated to the self when people engage in a similarity testing process. For instance, this model predicts that people are more likely to engage in similarity testing when they compare themselves to moderately high, rather than extremely high, standards of comparison (Mussweiler et al., 2004). This model can contribute to explain why we failed to find significant effects of subliminal comparison with average size models in Experiment 1. However, our results are not entirely consistent with this model, as we found no assimilation effect in this experiment (self-evaluation was not more positive in the average size models than in the no-models condition). Thus, this model is interesting but it does not seem to account for our findings. More generally, it does not clearly specify the conditions under which individuals are supposed to engage in a similarity or dissimilarity testing process. Thus, in its actual formulation this model does not allow making clear predictions about when contrast and assimilation effects are most likely to occur. In the same way, research on the active-self model (Wheeler, DeMarree, & Petty, 2007) suggests that subliminal primes are more likely to lead to assimilation effects in individuals low in self-monitoring. However, we assessed self-monitoring in the two experiments reported here, but failed to find any evidence that it moderated subliminal priming effects. Thus, further research is needed to better understand when and why subliminal primes lead to assimilation or contrast effects. Social comparison research has not made much progress in resolving this issue over the last decade.

A third limitation is that our experiments were underpowered to detect significant moderations. Previous research has clearly shown that women who are the most vulnerable (e.g., body dissatisfied women) are also more prone to the negative consequences of thin ideal exposures (e.g., Ferguson, 2013). In the present experiments, we found no evidence of moderation. It remains unclear why we failed to find support for moderation but one possible explanation is that our experiments did not have enough statistical power to detect moderation effects. Indeed, our sample size was based on power analyses designed to detect condition effects only. Thus, further research, with even larger samples, might be needed to detect a significant moderation by body

dissatisfaction. In the same way, because our samples included only young women from a western industrialized country, it remains unknown whether parallel effects could be found among men and in other, non-western countries. Research indicates that gender and cultural differences influence social comparison processes (Guimond & Chatard, 2013; Guimond et al., 2007). Thus, further research is needed to explore the generality of the findings reported here.

A final limitation rests in the difficulties involved in assessing awareness of stimuli in research that presents stimuli subliminally (Newell & Shanks, 2014). To show that stimuli can be presented without awareness (“too fast to be detected”), researchers have typically used binary forced-choice tasks, in which people are asked to indicate whether they can see briefly presented masked stimuli. If participants fail to discriminate the stimuli above chance level, researchers conclude that the primes are processed non-consciously. One problem with this approach is that it relies on testing the null hypothesis to establish that stimuli are undetectable (Rouder, Morey, Speckman, & Pratte, 2007; Sand & Nilsson, 2016). Proving the null hypothesis is difficult, especially with small samples and conventional analyses. As a matter of fact, some researchers have recently advocated for another approach that combines high-powered samples and the computation of Bayesian factors (Rouder et al., 2007; Sand & Nilsson, 2016). As Newell & Shanks (2014, p. 14) noticed, these “Recent methodological advances (e.g., Rouder et al., 2007) offer the promise of more clear-cut tests of subliminal perception in the future.” In the present experiments, we relied precisely on this recommended approach. All participants completed an objective awareness check at the end of the experiment.



Example of thin-ideal (left) and average-size (right) media images used in Experiment 1.

Bayesian analyses indicated that, given the data, the null hypothesis that subliminal primes were undetectable was much more likely than the hypothesis that they were detectable. Admittedly, it is difficult to be completely sure that participants are unaware of stimuli that are presented subliminally. However, we have specifically addressed this question experimentally to increase our confidence that this is the case, and all available indicators are indeed consistent with this idea.

4.3. Conclusion

The clear take-home message from this research is that social comparison takes place in the absence of awareness and affects explicit self-evaluation. This research provides converging evidence for the hypothesis that social comparison is an automatic process (Gilbert et al., 1995). One limitation that dampens the strength of previous research is that it is very difficult to compare the results of different

experiments because virtually all have used a different procedure. This is a major problem in social psychology when researchers attempt to independently replicate the effects of others (Hagger et al., 2016). Progress may now come from efforts to replicate existing research using standardized procedures. A standardized procedure (replication script), publicly available on the Open Science Framework, can be used to replicate our Experiment 2 in < 10 min. We are confident that our findings can be easily replicated, at least when using the exact same materials and procedure (see the Open practices section). During the peer-reviewing process, we conducted a pre-registered exact replication of Experiment 2 (see Experiment 3, available at <https://osf.io/43uwf/>). In this replication, we found similar findings than those reported here (for a summary of the main findings, see <https://osf.io/tqzdz/>). Our hope is that other researchers will also try to replicate the present findings, and that this will contribute to build a reliable and cumulative science on social comparison.

Open practices

All data and materials have been made publicly available via Open Science Framework and can be accessed at <https://osf.io/gjc5b/>. Experiment 1 was preregistered at <https://osf.io/gjc5b/registrations/>. Experiment 3 was preregistered at <https://osf.io/43uwf/registrations/>. The complete Open Practices Disclosure for this article can be found at <https://osf.io/v5x9s/>.

Appendix A

Appendix B. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jesp.2017.05.005>.

References

- Alicke, M. D. (2007). In defense of social comparison. *Revue Internationale de Psychologie Sociale*, 20, 11–29. Retrieved from <http://www.cairn.info/revue-internationale-de-psychologie-sociale-2007-1-page-11.htm>.
- Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman, & J. A. Bargh (Eds.), *Unintended thought* (pp. 3–51). New York: Guilford Press.
- Bargh, J. A. (1994). The four horsemen of automaticity. In R. S. Wyer, & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1–40). Hillsdale, NJ: Erlbaum.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230–244. <http://dx.doi.org/10.1037/0022-3514.71.2.230>.

- Blanton, H., & Stapel, D. A. (2008). Unconscious and spontaneous and... complex: The three selves model of social comparison assimilation and contrast. *Journal of Personality and Social Psychology*, 94, 1018–1032. <http://dx.doi.org/10.1037/0022-3514.94.6.1018>.
- Bocage-Barthélémy, Y., Chatard, A., Billieux, J., Daveau, E., Jafaari, N., & Selimbegović, L. (2017). *Automatic social comparison: Cognitive load facilitates an increase in negative thought accessibility after thin ideal exposure among women*. (Manuscript submitted for publication).
- Brown, A., & Dittmar, H. (2005). Think “thin” and feel bad: The role of appearance schema activation, attention level, and thin-ideal internalization for young women's responses to ultra-thin media ideals. *Journal of Social and Clinical Psychology*, 24, 1088–1113. <http://dx.doi.org/10.1521/jscp.2005.24.8.1088>.
- Callaway, E. (2011). Report finds massive fraud at Dutch universities: Investigation claims dozens of social-psychology papers contain faked data. *Nature*, 479, 15. <http://dx.doi.org/10.1038/479015a>.
- Chatard, A., Hirschberger, G., Faure, R., Champeaux, G., & Mermillod, M. (2017). *Liberty, equality, prejudice: Evidence that automatic activation of national identity increases implicit prejudice*. (Manuscript submitted for publication).
- Chatard, A., & Selimbegović, L. (2011). When self-destructive thoughts flash through the mind: Failure to meet standards affects the accessibility of suicide-related thoughts. *Journal of Personality and Social Psychology*, 100, 587–605. <http://dx.doi.org/10.1037/a0022461>.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159. <http://dx.doi.org/10.1037/0033-2909.112.1.155>.
- Dany, L., & Urdapilleta, I. (2012). Validation of a French measure of body comparison: The physical appearance comparison scale. *Revue Internationale de Psychologie Sociale*, 25, 97–112.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5–18. <http://dx.doi.org/10.1037/0022-3514.56.1.5>.
- Dittmar, H., & Howard, S. (2004). Thin-ideal internalization and social comparison tendency as moderators of media models' impact on women's body-focused anxiety. *Journal of Social and Clinical Psychology*, 23, 768–791. <http://dx.doi.org/10.1521/jscp.23.6.768.54799>.
- Ferguson, C. J. (2013). In the eye of the beholder: Thin-ideal media affects some, but not most, viewers in a meta-analytic review of body dissatisfaction in women and men. *Psychology of Popular Media Culture*, 2, 20–37. <http://dx.doi.org/10.1037/a0030766>.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140. <http://dx.doi.org/10.1177/001872675400700202>.
- Francis, G. (2012). The psychology of replication and replication in psychology. *Perspectives on Psychological Science*, 7(6), 585–594.
- Funder, D. C., Levine, J. M., Mackie, D. M., Morf, C. C., Sansone, C., Vazire, S., & West, S. G. (2014). Improving the dependability of research in personality and social psychology. Recommendations for research and educational practice. *Personality and Social Psychology Review*, 18, 3–12. <http://dx.doi.org/10.1177/1088868313507536>.
- Gana, K., & Brechenmacher, N. (2001). Structure latente et validité de la version française du self-monitoring scale: échelle de monitoring de soi/Latent, structure and validity of the French version of the self-monitoring scale. *L'Année Psychologique*, 101, 393–420. <http://dx.doi.org/10.3406/psy.2001.1338>.
- Garner, D. M., Olmstead, M. P., & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders*, 2, 15–34. [http://dx.doi.org/10.1002/1098-108X\(198321\)2:2<15::AID-EAT2260020203>3.0.CO;2-6](http://dx.doi.org/10.1002/1098-108X(198321)2:2<15::AID-EAT2260020203>3.0.CO;2-6).
- Gibbons, F. X., & Gerrard, M. (1989). Effects of upward and downward social comparison on mood states. *Journal of Social and Clinical Psychology*, 8, 14–31. <http://dx.doi.org/10.1521/jscp.1989.8.1.14>.
- Gilbert, D. T., Giesler, R. B., & Morris, K. A. (1995). When comparisons arise. *Journal of Personality and Social Psychology*, 69, 227–236. <http://dx.doi.org/10.1037/0022-3514.69.2.227>.
- Grabe, S., Ward, L. M., & Hyde, J. S. (2008). The role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies. *Psychological Bulletin*, 134, 460–476. <http://dx.doi.org/10.1037/0033-2909.134.3.460>.
- Groesz, L. M., Levine, M. P., & Murnen, S. K. (2002). The effect of experimental presentation of thin media images on body satisfaction: A meta-analytic review. *International Journal of Eating Disorders*, 31, 1–16. <http://dx.doi.org/10.1002/eat.10005>.
- Guimond, S. (2006). *Social comparison and social psychology: Understanding cognition, intergroup relations, and culture*. New York: Cambridge University Press.
- Guimond, S., Branscombe, N. R., Brunot, S., Buunk, A. P., Chatard, A., Désert, M., ... Yzerbyt, V. (2007). Culture, gender, and the self: Variations and impact of social comparison processes. *Journal of Personality and Social Psychology*, 92, 1118–1134. <http://dx.doi.org/10.1037/0022-3514.92.6.1118>.
- Guimond, S., & Chatard, A. (2013). Basic principles of social comparison: Does gender matter? In Z. Krizan, & F. X. Gibbons (Eds.), *Communal functions of social comparison* (pp. 205–229). Cambridge: Cambridge University Press.
- Guimond, S., Chatard, A., Martinot, D., Crisp, R. J., & Redersdorff, S. (2006). Social comparison, self-stereotyping, and gender differences in self-construals. *Journal of Personality and Social Psychology*, 90, 221–242. <http://dx.doi.org/10.1037/0022-3514.90.2.221>.
- Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., ... Zwienerberg, M. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, 11, 546–573. <http://dx.doi.org/10.1177/1745691616652873>.
- Harris, C. V., Bradlyn, A. S., Coffman, J., Gunel, E., & Cottrell, L. (2008). BMI-based body size guides for women and men: Development and validation of a novel pictorial method to assess weight-related concepts. *International Journal of Obesity*, 32, 336–342. <http://dx.doi.org/10.1038/sj.ijo.0803704>.
- Heinberg, L. J., Thompson, J. K., & Stormer, S. (1995). Development and validation of the sociocultural attitudes towards appearance questionnaire. *International Journal of Eating Disorders*, 17, 81–89. [http://dx.doi.org/10.1002/1098-108X\(199501\)17:1<81::AID-EAT2260170111>3.0.CO;2-Y](http://dx.doi.org/10.1002/1098-108X(199501)17:1<81::AID-EAT2260170111>3.0.CO;2-Y).
- Henderson-King, E., & Henderson-King, D. (1997). Media effects on women's body esteem: Social and individual difference factors. *Journal of Applied Social Psychology*, 27, 399–417. <http://dx.doi.org/10.1111/j.1559-1816.1997.tb00638.x>.
- Jansen, A., & de Vries, M. (2002). Pre-attentive exposure to the thin female beauty ideal does not affect women's mood, self-esteem and eating behaviour. *European Eating Disorders Review*, 10, 208–217. <http://dx.doi.org/10.1002/erv.430>.
- Jeffreys, H. (1961). *Theory of probability*. Oxford: Oxford University Press.
- Kawakami, K., Dovidio, J. F., & Dijksterhuis, A. (2003). Effect of social category priming on personal attitudes. *Psychological Science*, 14, 315–319. <http://dx.doi.org/10.1111/1467-9280.14451>.
- Kim, J. G., Biederman, I., Lescroart, M. D., & Hayworth, K. J. (2009). Adaptation to objects in the lateral occipital complex (LOC): Shape or semantics? *Vision Research*, 49, 2297–2305. <http://dx.doi.org/10.1016/j.visres.2009.06.020>.
- Mills, J. S., Polivy, J., Herman, C. P., & Tiggenmann, M. (2002). Effects of exposure to thin media images: Evidence of self-enhancement among restrained eaters. *Personality and Social Psychology Bulletin*, 28, 1687–1699.
- Morse, S., & Gergen, K. J. (1970). Social comparison, self-consistency, and the concept of self. *Journal of Personality and Social Psychology*, 16(1), 148–156. <http://dx.doi.org/10.1037/h0029862>.
- Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, 64, 723–739. <http://dx.doi.org/10.1037/0022-3514.64.5.723>.
- Mussweiler, T. (2003). Comparison processes in social judgment. Mechanisms and consequences. *Psychological Review*, 110, 472–489. <http://dx.doi.org/10.1037/0033-295X.110.3.472>.
- Mussweiler, T., & Bodenhausen, G. V. (2002). I know you are, but what am I? Self-evaluative consequences of judging in-group and out-group members. *Journal of Personality and Social Psychology*, 82, 19–32. <http://dx.doi.org/10.1037/0022-3514.82.1.19>.
- Mussweiler, T., Rüter, K., & Epstude, K. (2004). The man who wasn't there: Subliminal social comparison standards influence self-evaluation. *Journal of Experimental Social Psychology*, 40, 689–696. <http://dx.doi.org/10.1016/j.jesp.2004.01.004>.
- Newell, B. R., & Shanks, D. R. (2014). Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences*, 37, 1–19. <http://dx.doi.org/10.1017/S0140525X12003214>.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349, aac4716. <http://dx.doi.org/10.1126/science.aac4716>.
- Pavlo, M. (2006). How drive for thinness moderates the influence of media on female body image. *Journal of Undergraduate Psychological Research*, 1, 22–26.
- Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, 162, 8–13. <http://dx.doi.org/10.1016/j.jneumeth.2006.11.017>.
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2, 10. <http://dx.doi.org/10.3389/neuro.11.010.2008>.
- Plant, R. R. (2016). A reminder on millisecond timing accuracy and potential replication failure in computer-based psychology experiments: An open letter. *Behavior Research Methods*, 48, 408–411. <http://dx.doi.org/10.3758/s13428-015-0577-0>.
- Posavac, H. D., Posavac, S. S., & Posavac, E. J. (1998). Exposure to media images of female attractiveness and concern with body weight among young women. *Sex Roles*, 38, 187–201. <http://dx.doi.org/10.1023/A:1018729015490>.
- Reed, D. L., Thompson, J. K., Brannick, M. T., & Sacco, W. P. (1991). Development and validation of the physical appearance state and trait anxiety scale (PASTAS). *Journal of Anxiety Disorders*, 5, 323–332. [http://dx.doi.org/10.1016/0887-6185\(91\)90032-O](http://dx.doi.org/10.1016/0887-6185(91)90032-O).
- Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7, 331–363. <http://dx.doi.org/10.1037/1089-2680.7.4.331>.
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Pratte, M. S. (2007). Detecting chance: A solution to the null sensitivity problem in subliminal priming. *Psychonomic Bulletin & Review*, 14, 597–605. <http://dx.doi.org/10.3758/BF03196808>.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16, 225–237. <http://dx.doi.org/10.3758/PBR.16.2.225>.
- Rousseau, A., Knotter, R. M., Barbe, R. M., Raich, R. M., & Chabrol, H. (2005). Étude de validation de la version française du body shape questionnaire. *L'Encéphale*, 31, 162–173.
- Rousseau, A., Valls, M., & Chabrol, H. (2010). The sociocultural attitudes towards appearance scale-3 (SATAQ-3): étude de validation de la version française. *L'Encéphale*, 36, 270–276. <http://dx.doi.org/10.1016/j.encep.2008.03.013>.
- Sand, A., & Nilsson, M. E. (2016). Subliminal or not? Comparing null-hypothesis and Bayesian methods for testing subliminal priming. *Consciousness and Cognition*, 44, 29–40. <http://dx.doi.org/10.1016/j.concog.2016.06.012>.
- Schimmack, U. (2012). The ironic effect of significant results on the credibility of multiple-study articles. *Psychological Methods*, 17, 551–566. <http://dx.doi.org/10.1037/a0029487>.
- Schimmack, U. Quantifying statistical research integrity: The replicability index. (2014). Retrieve from http://www.r-index.org/uploads/3/5/6/7/3567479/introduction_to_the_r-index_14-12-01.pdf (Unpublished manuscript).
- Simons, D. J. (2014). The value of direct replication. *Perspectives on Psychological Science*, 9, 76–80. <http://dx.doi.org/10.1177/1745691613514755>.
- Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, 26, 559–569. <http://dx.doi.org/10.1177/>

- 0956797614567341.
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, 143, 534–547. <http://dx.doi.org/10.1037/a0033242>.
- Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, 30, 526–537. <http://dx.doi.org/10.1037/h0037039>.
- Stapel, D. A., & Blanton, H. (2004). From seeing to being: Subliminal social comparisons affect implicit and explicit self-evaluations. *Journal of Personality and Social Psychology*, 87, 468–481. <http://dx.doi.org/10.1037/0022-3514.87.4.468>.
- Stojanoski, B., & Cusack, R. (2014). Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. *Journal of Vision*, 14(12), 6. <http://dx.doi.org/10.1167/14.12.6>.
- Suls, J., Martin, R., & Wheeler, L. (2002). Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science*, 11, 159–163.
- Taylor, S. E., & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. *Psychological Review*, 96, 569–575. <http://dx.doi.org/10.1037/0033-295X.96.4.569>.
- Thompson, J. K., Heinberg, L., & Tantleff-Dunn, S. (1991). The physical appearance comparison scale (PACS). *The Behavior Therapist*, 14, 174.
- Vogels, R. (1999). Effect of image scrambling on inferior temporal cortical responses. *Neuroreport*, 10(9), 1811–1816. <http://dx.doi.org/10.1097/00001756-199906230-00002>.
- Want, S. C. (2009). Meta-analytic moderators of experimental exposure to media portrayals of women on female appearance satisfaction: Social comparisons as automatic processes. *Body Image*, 6, 257–269. <http://dx.doi.org/10.1016/j.bodyim.2009.07.008>.
- Want, S. C., Botres, A., Vahedi, Z., & Middleton, J. A. (2015). On the cognitive (in) efficiency of social comparisons with media images. *Sex Roles*, 73, 519–532. <http://dx.doi.org/10.1007/s11199-015-0538-1>.
- Want, S. C., & Saiphoo, A. (2017). Social comparisons with media images are cognitively inefficient even for women who say they feel pressure from the media. *Body Image*, 20, 1–6. <http://dx.doi.org/10.1016/j.bodyim.2016.10.009>.
- Wetzels, R., Raaijmakers, J. G., Jakab, E., & Wagenmakers, E. J. (2009). How to quantify support for and against the null hypothesis: A flexible WinBUGS implementation of a default Bayesian *t* test. *Psychonomic Bulletin & Review*, 16, 752–760. <http://dx.doi.org/10.3758/PBR.16.4.752>.
- Wheeler, S. C., DeMarree, K. G., & Petty, R. E. (2007). Understanding the role of the self in prime-to-behavior effects: The active-self account. *Personality and Social Psychology Review*, 11, 234–261. <http://dx.doi.org/10.1177/1088868307302223>.
- Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin*, 127, 797–826. <http://dx.doi.org/10.1037/0033-2909.127.6.797>.