REPLICATION CORNER

How expressing one's likes and dislikes affects enjoyment: a replication

Julia Sophia Wittich[1] · Jan R. Landwehr[2] · Daniel Wentzel[1] ⓘ

Abstract

A recent article by He, Melumad, and Pham (Journal of Consumer Research 46(3):545–563, 2019) showed that consumers experienced greater enjoyment when they were asked to perform an evaluative task relative to a non-evaluative task. In this research, we intend to replicate this finding and to examine to what extent the strength of the effect is contingent on the sample employed. Using a sample from Amazon MTurk, study 1 replicates the indirect effect but not the total experimental effect. Study 2 draws on a sample that is unacquainted with the experimental paradigm and replicates the indirect as well as the total experimental effect. These findings suggest that the sampling population of MTurk may have become desensitized to the basic experimental paradigm.

1 Introduction

Judging and evaluating the environment is part of our daily lives. We listen to a pop song and share our opinion, we watch a YouTube video and leave a comment, and we read and "like" the latest post of an influencer on Instagram. Standard consumer behavior theory assumes that such evaluations mainly serve an instrumental purpose. That is, forming an evaluation of a market-related stimulus (e.g., a

---

✉ Daniel Wentzel
 wentzel@time.rwth-aachen.de

Julia Sophia Wittich
 wittich@time.rwth-aachen.de

Jan R. Landwehr
 landwehr@wiwi.uni-frankfurt.de

[1] School of Business and Economics, RWTH Aachen University, Templergraben 55, 52062 Aachen, Germany

[2] Faculty of Economics and Business, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 4, 60629 Frankfurt am Main, Germany

Ⓓ Springer

product, service, brand, or ad) may inform future buying or consumption decisions. In many cases, however, consumers form opinions even if doing so does not serve an explicit instrumental purpose. For instance, why do eleven million users "like" a post of Cristiano Ronaldo taking a walk in the snow or five million "like" a picture of Selena Gomez wearing a Puma tracksuit when these evaluations are not tied to any particular purchase decision?

In a recent study, He et al. (2019) found that consumers derive an inherent pleasure from expressing their likes and dislikes. A total of seven experiments showed that consumers who make evaluation-based judgments (i.e., stating whether they like a particular product) experience higher enjoyment during this process compared to consumers making non-evaluation-based judgments (i.e., rating a product on a particular attribute).

The authors found that this effect occurs because evaluative judgments allow consumers to express their selves. As such, self-expression, the act of disclosing one's identity, is a fundamental human need that can be accomplished through many different acts, including one's choices, preferences, and attitudes (Katz, 1960). Indeed, many purchase decisions are driven by consumers' desire to build and express a version of their self (Belk, 1988; Chernev et al., 2011). Moreover, expressing one's self is pleasurable, whereas suppressing the disclosure of self-relevant information leads to negative feelings (Pennebaker et al., 1988). In light of these findings, He et al. (2019) argue that consumers may consider evaluating products as inherently rewarding. While many such evaluations may not serve an explicit instrumental purpose, they allow consumers to express their identity. Put differently, by communicating what they like and dislike, consumers can disclose who they are.

While the results of He et al. (2019) are comprehensive, many of the effect sizes are comparatively small (η^2: 0.006–0.069). As six out of seven studies were conducted with samples of Amazon MTurk, the authors point to the possibility that the sampled population of MTurk participants may have become desensitized to the basic experimental paradigm. That is, assuming that at least some of the MTurk workers participated in more than one study, the effect of the experimental manipulation (i.e., providing liking vs. non-liking judgments) may have become less pronounced over time.

On the one hand, this may be considered a methodological artifact that may generally affect research projects that are based on multiple studies drawing on the same sampling population. On the other hand, a potential desensitization of participants may directly affect the nature of the underlying effect. That is, the account provided by He et al. (2019) argues that consumers inherently enjoy expressing their likes because they consider these tasks as expressions of their identity. If, however, this effect is weakened, the more often consumers are asked to express their likes (i.e., if they become desensitized over time), then this raises the question if the effect only materializes for specific consumer groups and/or specific consumption settings.

Against this background, the aim of this study is to examine the robustness of the effects of He et al. (2019) with the use of different samples. Study 1 relies on a sample of MTurk participants and replicates the experimental effect on the mediator and the indirect effect reported by He et al. (2019). The total effect, however, does not replicate. Study 2 draws on a student sample from a German university and

replicates the total as well as the indirect effect. The data and statistical code of both studies are available on the Open Science Framework (OSF): https://osf.io/qu7zf/.[1]

2 Study 1

Study 1 intends to replicate the key finding of He et al. (2019) — namely, that evaluative judgments lead to greater task enjoyment than non-evaluative judgments. To this end, we focused on a specific contrast of the second study of He et al. (2019). The initial version of this study is reported as study 2B in the online appendix and includes several contrasts between evaluative vs. non-evaluative judgments (He et al., 2019, p. 6). The strongest effect (see p. 2 of their Web Appendix) is revealed for the contrast between liking judgments (i.e., evaluative) and casualness judgments (i.e., non-evaluative), which is why we decided to focus on the comparison of these two conditions. We did not directly replicate the original study but used a new set of t-shirts with a more common design as stimulus material. In line with He et al. (2019), we examined whether evaluative judgments lead to higher task enjoyment than non-evaluative judgments and whether this effect is mediated by self-expressiveness.

2.1 Design, procedure, and participants

The study protocol was preregistered on AsPredicted: https://aspredicted.org/m26ya.pdf. The study employs a one-factorial between-subjects design (type of evaluation: evaluative vs. non-evaluative). All items, unless otherwise stated, were taken from He et al. (2019). In the evaluative condition, participants were asked to express their liking of ten t-shirts of a large online retailer, whereas in the non-evaluative condition, participants evaluated the casualness of the same ten t-shirts. Afterward, participants responded to two items (1 = "I did not enjoy this task at all"; 7 = "I enjoyed this task very much" and 1 = "I feel it was not fun at all"; 7 = "I feel it was very fun"), which were averaged ($\alpha = 0.82$) to form the dependent variable task enjoyment. In addition, participants reported the degree of self-expressiveness (1 = "I expressed very little about myself" to 7 = "I expressed a lot about myself"). Participants also reported their task involvement, answered some personality measures collected for exploratory purposes, and provided their demographics.

In line with He et al. (2019, study 2), we aimed for 500 participants per experimental condition and recruited a total of 1046 participants ($M_{age} = 36.06$, $SD_{age} = 11.16$, 35.9% female) from Amazon MTurk. In contrast to our preregistration, we did not exclude any participants from our analyses because the pattern of results does not change when participants with a low level of attention are excluded (see OSF for more details).

---

[1] The OSF project also includes the data and code of a third study which examined a moderation of the core effect and which is not reported here.

2.2 Results and discussion

The results of an ANOVA revealed no significant differences in terms of task enjoyment between the evaluative and non-evaluative conditions ($F(1, 1044) = 0.99$, $p = 0.320$; $M_{eval} = 5.59$, SD $= 1.25$; $M_{non-eval} = 5.51$, SD $= 1.26$, $\eta^2 = 0.001$). However, there were significant differences in terms of self-expressiveness across the experimental groups ($F(1, 1044) = 4.79$, $p = 0.029$; $M_{eval} = 5.37$, SD $= 1.40$; $M_{non-eval} = 5.17$, SD $= 1.61$, $\eta^2 = 0.005$). A process analysis using model 4 of the PROCESS macro with 5000 bootstrapped samples (Hayes, 2013) revealed an indirect effect of evaluation type (effect-coded: $1 =$ evaluative task; $-1 =$ non-evaluative task) on task enjoyment through perceived self-expressiveness ($b = 0.051$, 95% CI [0.0064, 0.0982]). Hence, the experimental effect on the mediator and the indirect effect replicate the results of He et al. (2019), while the total effect does not replicate.

3 Study 2

He et al. (2019, p. 6) argue that the MTurk population may be desensitized to the employed experimental paradigm, which may explain our failed replication of the total effect in study 1 and the rather low effect size. Accordingly, we decided to run another replication with a different sample of participants (i.e., a student sample) that is unfamiliar with the experimental paradigm and potentially more involved when participating in the study. For a fresh sample of participants, He et al. (2019) report an Eta-squared of 0.044 (see p. 2 Study 2B of their Web Appendix) for their experimental effect, which indicates that a sample of $N = 174$ would be sufficient to detect an effect with a power of 0.80. Based on this analysis, we recruited $N = 204$ German university students ($M_{age} = 20.85$, SD$_{age} = 3.13$, 42.2% female) from a marketing lecture who participated voluntarily. Apart from the different samples, study 2 was identical to study 1. That is, we relied on the same stimuli, experimental procedure, and measures as in study 1.

3.1 Results and discussion

Consistent with He et al. (2019), the results of an ANOVA show that people enjoy an evaluative task more than a non-evaluative task ($F(1, 202) = 16.15$, $p < 0.001$; $M_{eval} = 4.52$, SD $= 1.33$; $M_{non-eval} = 3.72$, SD $= 1.51$, $\eta^2 = 0.074$). Furthermore, participants reported higher self-expressiveness in the evaluative group than in the non-evaluative group ($F(1, 202) = 4.239$, $p = 0.041$; $M_{eval} = 4.09$, SD $= 1.86$; $M_{non-eval} = 3.55$, SD $= 1.86$, $\eta^2 = 0.021$). A process analysis using model 4 of the PROCESS macro with 5000 bootstrapped samples (Hayes, 2013) revealed that perceived self-expressiveness mediated the effect of evaluation type (effect-coded: $1 =$ evaluative task; $-1 =$ non-evaluative task) on task enjoyment ($b = 0.093$, 95% CI [0.0043, 0.1903]).

In sum, study 2 successfully replicates the findings of He et al. (2019) by showing that an evaluative task triggers higher task enjoyment than a non-evaluative task and that this effect is mediated by increased self-expressiveness. Notably, when relying on a student sample, we observe a stronger effect size for the total effect of the experimental manipulation on task enjoyment ($\eta^2 = 0.074$) than He et al. (2019) in their studies 2 ($\eta^2 = 0.015$), 2B ($\eta^2 = 0.044$), and 2C ($\eta^2 = 0.063$). Finally, a post hoc power analysis shows that our replication study achieved a power of 0.98, confirming the adequacy of our sample size.

4 Meta-analytic comparison across studies

Given that we were interested in examining if the strength of the observed effect is contingent on the specific sample employed, we conducted a meta-analytic comparison covering studies 1 and 2 as well as study 2 from He et al. (2019). This analysis allowed us to test whether there is an overall effect across all studies and if the strength of this effect differs between the studies. We entered the biased-corrected standardized mean differences and their sampling variances in the R function "rma" from the "metafor"-package (Viechtbauer, 2010). This analysis yielded a significant meta-analytical estimate for the main effect of task type (evaluative vs. non-evaluative) on task enjoyment ($b = 0.278$; SE $= 0.136$; $z = 2.038$; $p = 0.042$). Moreover, the meta-analytical estimate for the heterogeneity between the studies was also significant ($p = 0.002$), indicating that there is more heterogeneity between the studies than would be expected based on sampling variability alone (Viechtbauer, 2010).

5 General discussion

He et al. (2019) demonstrated that consumers derive an inherent pleasure from expressing their likes and dislikes. This finding potentially holds great relevance for consumer research as it may not only help explain why "liking" has become such a pervasive phenomenon but may also shed new light on how consumers engage in identity-building behaviors in contemporary environments. Hence, we aimed to replicate the key finding from He et al. (2019). Study 1 drew on a similar sample from MTurk as the original studies. While the indirect effect replicated, the total experimental effect did not replicate. Study 2 drew on a sample that was unacquainted with the experimental paradigm (German university students) and replicated the indirect as well as the total experimental effect. Moreover, while a post hoc meta-analytic comparison provides support for an overall effect, it also reveals that there is substantial variability in the effect size across the three different samples. As such, the strength of the observed effect is contingent on the specific sample employed. Overall, these results are consistent with the notion that samples drawn from the MTurk population may have become desensitized to the experimental paradigm and that this desensitization may affect the strength of the experimental effect.

These findings have important methodological and substantial implications. From a methodological perspective, our findings suggest that the extent to which a "liking"

effect materializes is contingent on the specific sample, with samples that may have been exposed to similar studies in the past responding less strongly to the experimental manipulations. While such effects may be common to studies that draw on the same sampling population, a potential desensitization may also have more substantial implications. That is, He et al. (2019) argue that consumers consider expressing their preferences as an identity-relevant task and will actually welcome the opportunity to do so. In digital environments, however, consumers are constantly encouraged to articulate their likes. Globally, consumers spend roughly 2.5 h every day on social media apps (Statista, 2021), and a central feature of most apps consists in consuming and evaluating user- or firm-generated content.

If consumers indeed become desensitized to expressing their preferences, then this may call into question if and to what extent the effect will emerge in real environments. Over time, consumers may satiate from articulating their likes and may stop believing that doing so is a form of expressing their identity. For instance, while a "liking" effect may be observed when an app is initially launched and/or when users are new to the app, this effect may level out once users have become desensitized to expressing their likes. While these arguments have not been tested in our studies and are thus somewhat speculative, they suggest that more research is needed to fully understand the scope of the "liking" effect.

Declarations

Ethics approval Obtained.

Conflict of interest The authors declare no competing interests.

References

Belk, R. W. (1988). Possessions and the extended self. *Journal of Consumer Research, 15*(2), 139.

Chernev, A., Hamilton, R., & Gal, D. (2011). Competing for consumer identity: Limits to self-expression & the perils of lifestyle branding. *Journal of Marketing, 75*(3), 66–82.

Hayes, A. F. (2013). *An introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

He, D., Melumad, S., & Pham, M. T. (2019). The pleasure of assessing and expressing our likes and dislikes. *Journal of Consumer Research, 46*(3), 545–563.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly*, *24*(2, Special Issue: Attitude Change), 163–204.

Pennebaker, J. W., Kiecolt-Glaser, J. K., & Glaser, R. (1988). Disclosure of traumas and immune function: Health implications for psychotherapy. *Journal of Consulting and Clinical Psychology, 56*(2), 239–245.

Statista. (2021). Statista digital economy compass. https://www.statista.com/study/105653/digital-economy-compass/. Accessed 15 Jan 2023.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor. *Journal of Statistical Software, 36*, 1–48.