[Link to supplementary: [Purple [Group B] - Replication & Extension - Supplementary](https://docs.google.com/document/d/1OYgFI_E-ks4khwz2hQyg4v5FKgRf0RUroDk1uyBJVm4/edit)]

Newman et al. (2011):   
Replication and extension(s)

Mannix Chan  
Department of Psychology, University of Hong Kong, Hong Kong SAR  
[mannix42@connect.hku.hk](mailto:mannix42@connect.hku.hk) / [mannixchan9149@gmail.com](mailto:mannixchan9149@gmail.com)

Yaqi Jin  
Department of Psychology, University of Hong Kong, Hong Kong SAR  
[jinyaqi@connect.hku.hk](mailto:jinyaqi@connect.hku.hk) / [yvonne27hku@gmail.com](mailto:yvonne27hku@gmail.com)

[Name]  
Department of Psychology, University of Hong Kong, Hong Kong SAR  
[Email 1] / [Email 2]

\*Contributed equally, joint first author

^Corresponding author

Word: abstract – [XXX], manuscript - [XXXX]

[**Note**: journals defer on how they calculate manuscript word count. Typically, it refers to the text included in the main sections, without tables/figures/references/appendices, and without anything that comes before "introduction"]

## Author bios:

Mannix Chan and Yaqi Jin were students at the University of Hong Kong during the academic year 2021-2022.

[TA1 first name] [TA1 last name] were teaching assistants at the University of Hong Kong psychology department during the academic year [XXXX].

[Gilad Feldman is an assistant professor with the University of Hong Kong psychology department. His research focuses on judgment and decision-making.]

## Declaration of Conflict of Interest:

The author(s) declared no potential conflicts of interests with respect to the authorship and/orpublication of this article.

## Financial disclosure/funding:

The author(s) received no financial support for the research and/or authorship of this article.

## Authorship declaration:

[Can use the [Tenzing tool](https://martonbalazskovacs.shinyapps.io/tenzing/)]

Mannix Chan and Yaqi Jin conducted the replication as part of the PSYC2071 Judgments and decision making course.

[TA1 first name] [TA1 last name] guided and assisted the replication effort in the [Course name] [Course code].

[Gilad] was the course instructor for [Course name] [Course code] and led the replication efforts in the[se] course[s]. [Gilad] supervised each step in the project, conducted the pre-registrations, and ran data collection.

## Corresponding author

[Gilad Feldman, Department of Psychology, University of Hong Kong, Hong Kong SAR; [gfeldman@hku.hk](mailto:gfeldman@hku.hk) ; 0000-0003-2812-6599]

## Rights:

CC BY or equivalent licence is applied to the AAM arising from this submission. ([clarification](https://bit.ly/rrs-primer))

## **Full details of all authors**:

| First name | Last name | Researchgate profile | [ORCID](https://orcid.org/register) | OSF profile | Institutional email | Personal email | ID |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Mannix | Chan | <https://www.researchgate.net/profile/Mannix-Chan-2> | https://orcid.org/  0000-0001-9699-9522 | osf.io/6t7yc | [mannix42@connect.hku.hk](mailto:mannix42@connect.hku.hk) | [mannixchan9149@gmail.com](mailto:mannixchan9149@gmail.com) | 3035692589 |
| Yaqi | Jin | <https://www.researchgate.net/profile/Yaqi-Jin-4> | https://orcid.org/0000-0001-6745-7599 | osf.io/4xqr6 | [jinyaqi@connet.hku.hk](mailto:jinyaqi@connet.hku.hk) | [yvonne27hku@gmail.com](mailto:yvonne27hku@gmail.com) | 3035535339 |

\*Note: ORCIDs are now the de-facto standard for ensuring the proper credit across publications. Make sure that your ORCID has up-to-date information where you can be reached long term and has the name you wish to use in academia and includes a personal email you can receive updates and credit after graduation.

## Contributor Roles Taxonomy

In the table below, employ CRediT (Contributor Roles Taxonomy) to identify the contribution and roles played by the contributors in the current replication effort. Please refer to the url (<https://www.casrai.org/credit.html> ) on details and definitions of each of the roles listed below.

[Alternatively, use [Tenzing Documenting contributorship with CRediT](https://rollercoaster.shinyapps.io/tenzing/) (explained [here](https://medium.com/@ceptional/announcing-tenzing-ceca6789d88c))]

Table [X]

| **Role** | **Lead author (you?)** | **Corresponding author** | **Mannix Chan** | **Yaqi Jin** | **Other authors ( Teaching assistants)** |
| --- | --- | --- | --- | --- | --- |
| Conceptualization |  | X |  |  |  |
| Pre-registration |  | X | X | X |  |
| Data curation |  | X |  |  |  |
| Formal analysis | X | X | X | X |  |
| Funding acquisition |  | X |  |  |  |
| Investigation | X | X |  |  | X |
| Pre-registration peer review / verification |  |  | X | X | X |
| Data analysis peer review / verification | X |  |  |  | X |
| Methodology |  | X | X | X |  |
| Project administration |  | X |  |  | X |
| Resources |  | X |  |  |  |
| Software | X | X |  |  |  |
| Supervision |  | X |  |  | X |
| Validation | X | X |  |  |  |
| Visualization | X | X |  | X |  |
| Writing-original draft | X | X | X | X |  |
| Writing-review and editing | X | X |  |  |  |

**Important links and information**

Citation of the target research article:

Newman, G. E., Diesendruck, G., & Bloom, P. (2011). *Celebrity contagion and the value of objects. Journal of Consumer Research, 38*(2), 215-228.

Link to the target research article:

<https://doi.org/10.1086/658999>

**Links to project files**

(test ALL these using Incognito mode to make sure it's open for external readers!)

| **Content** | **Link** |
| --- | --- |
| Cloud folder for datasets and final analysis code | <enter link here> |
| pre-registration on the Open Science Framework | <enter link open for reviewers> |
| Cloud folder for pre-registration + Qualtrics + random dataset + prereg analysis code: | <https://drive.google.com/drive/folders/1NVtpRlVVb66O0iKdfyyBackYBKYgg7JC?usp=sharing> |
| Cloud folder for presentation (PPT and PDF) | <enter link here> |
| Cloud folder for peer reviews | <enter link here> |

**Additional information**

Add whatever other information you think is important about the project.

We thank Dr. George Newman for graciously providing us with materials used in the original experiments and for reviewing the Qualtrics survey.

The current replication is part of the larger ‘mass pre-registered replications in judgment and decision-making’ project. The project aims to revisit well known research findings in the area of judgment and decision making (JDM) and investigate the replicability of these findings. As part of the initiative the students engage in pre-registered replications to examine the well-known findings as part of regular one-semester coursework.

# Before you submit

[remove this section when submitting]

Export the document as a file and:

1. Please run preprint health check on your exported file: <https://app.scholarcy.com/preprint-check.html> (it will extract the references and find the DOIs for you]
2. Verify citations with [reciteworks.com](https://reciteworks.com/) , add DOIs with [crossref](https://doi.crossref.org/simpleTextQuery).
3. Run [statcheck](https://www.google.com/search?q=statcheck).
4. See [RRR assessment template](http://mgto.org/RRRassessment) and examine what can be used from that to improve on your submission.

# Abstract

**[IMPORTANT:   
Results were written using a randomized dataset produced by Qualtrics to simulate what these sections will look like after data collection. These will be updated following the data collection.]**

Contagion refers to the belief that physical objects can contain “remnants” of their previous owners. In a pre-registered experiment with an/a [sample, e.g. American online Amazon Mechanical Turk sample, British Prolific sample, HKU undergraduate sample] (*N* = XXX), we replicated Experiments 1 and 2 from Newman et al. (2011). [Brief description of **replication** findings with effects sizes and CIs of original and replication]. Thus, we found [weak to no / weak / medium / strong] empirical support for the hypotheses of the original article. Extending the replication, [Brief description of **extension** findings with effects sizes and CIs]. Supplementary materials, raw data and analysis files/code are available here: [insert OSF link].

*Keywords:* contagion, bias, judgment and decision making, registered replication, [important related keyword 1], [important related keyword 2]

# Stage 1 Snapshot

[Note: this is meant for the [PCI-RR scheduled track](https://osf.io/ecbqd/) ([info here](https://rr.peercommunityin.org/help/guide_for_authors#h_61998243643551613309672490)), and since 2021 is required for students’ reports. Please fill this best you can, included in the Part 1 submission.]

## ‎Research question(s) and/or theory

[Instructions: Briefly summarise the research question(s) that will be addressed, and ‎where relevant, the theoretical basis of the proposal.]

Direct replication and extensions of Newman et al. (2011); attempt to replicate the effects found, and investigating additional measures to exclude confounding factors

## Hypotheses

[Instructions: State all predictions of the study briefly. Can use the summary table from the introduction.]‎

Statistically significant effects found in Newman et al. (2011) will also be found in this study and vice versa. Original article analyses can be found in the Appendix.

Replication hypothesis for experiment 1:

* More contact was wanted with positive than negative celebrity items
* Contagion and market forces can predict valuation of celebrity items
  + Positive celebrity possessions value increased for increase in both contagion and market value
  + Negative celebrity possessions value increased for increase in market value and decreased in increase of market value

Original/replication hypothesis for experiment 2 can be found in table [X].

## Study design and methods

[Instructions: Summarise in broad terms the study design, including (as applicable), key ‎conditions and controls, data acquisition procedures, and variables.‎ Can use the table from the methods section based on our replication guide.]

Both experiments were conducted with an online survey.

Experiment 1:

| **IV1: Valence of named individual** [between subjects]  **IV2: Fame of named individual** [between subjects] | **IV1: Positive valence**  Participant requested to name an individual whom they consider to be “incredibly moral” | **IV1: Negative valence**  Participant requested to name an individual whom they consider to be “incredibly immoral” | **IV1: Mixed valence**  Participant requested to name an individual whom they consider to be “moral at times and immoral at other times” |
| --- | --- | --- | --- |
| **IV2: Celebrity**  Participant requested to name a celebrity or a public figure | **Dependent variables**  DV title: **Object valuation** (Primary dependent measure)  Specific DV items:  “How much would you like to own [this item]?”  “How likely would you be to purchase [this item] if it was for sale?”  “Is [this item] worth keeping, or would you throw it out?”  All of the above items rated on a scale of 1 (not at all) to 9 (very much so)  DV title: **Contagion**  Specific DV items:  “How much would you want to give this person a hug or shake their hand?”  “How much would you like to hold this item in your hands?”  All of the above items rated on a scale of 1 (not at all) to 9 (very much so)  DV title: **Market value**  Specific DV items:  “Are there some people who would pay money for this item?”  “Would other people be impressed if they found out that you owned this item?”  All of the above items rated on a scale of 1 (not at all) to 9 (very much so)  DV title: **Historical significance**  Specific DV item:  “Does this item have historical value (i.e. should it be put in a museum?)”, rated on a scale of 1 (not at all) to 9 (very much so)  DV title: **Liking**  Specific DV item:  “How much do you like this person?”, rated on a scale of 1 (extreme disliking) to 9 (extreme liking)  **Extension DV item**:  “How much would you like to meet this person?”, rated on a scale of 1 (not at all) to 9 (very much so)  Manipulation check:  “How famous is this person?”, rated on a scale of 1 (not at all famous) to 9 (extremely famous) | | |
| **IV2: Noncelebrity**  Participant requested to name a person whom they know personally |

Experiment 2:

|  | **IV2: Manipulation of physical contact**  Participants read a description where the degree of physical contact of the celebrity with a sweater was manipulated | | **IV2: Manipulation of market demand**  Participants read a description where the market demand for the celebrity’s sweater was manipulated | |
| --- | --- | --- | --- | --- |
| **IV3: Highlighted direction**  The celebrity’s degree of physical contact with the sweater was highlighted (“it was one of their favourite sweaters”) | **IV3: Decreased direction**  The celebrity’s degree of physical contact with the sweater was decreased (“they never […] even opened the box that it came in”) | **IV3: Highlighted direction**  The sweater’s market demand was highlighted (“there is a lot of demand for items owned by [the celebrity]”) | **IV3: Decereased direction**  The sweater’s market demand was decreased (“there is very little demand for items owned by [the celebrity]”) |
| **IV1: Positive celebrity**  Participant requested to name a celebrity whom they deeply admired | **Dependent variables**  DV title: **Willingness to purchase item**  Specific DV item:  “(Imagine that you have the opportunity to bid on a sweater that belonged to [the celebrity].) How willing are you to purchase this sweater compared to an identical used sweater (in the same condition) that was not owned by [the celebrity]?”, rated on a scale of 1 (much less likely to purchase) to 9 (much more likely to purchase)  DV title: **Pleasantness of wearing item**  Specific DV item:  “How pleasant would you find the experience of wearing [the celebrity]'s sweater?”, rated on a scale of 1 (extremely unpleasant) to 9 (extremely pleasant)  **Extension DV items**:  “Are there some people who would pay money for this item?”  “Would other people be impressed if they found out that you owned this item?”  All of the above items rated on a scale of 1 (not at all) to 9 (very much so)  Dependent variables and extension DV items are measured once before and once after the manipulation  Manipulation checks:  “Is [the celebrity] somebody you know personally?”  “Is [the celebrity] somebody that the average person has heard of?”  All of the above items rated on a binary measure | | | |
| **IV1: Negative celebrity**  Participant requested to name a celebrity whom they deeply despised |
| **IV individual differences predictor**: **Individual contagion sensitivity**  Specific items:  “Even if I were hungry, I would not drink a bowl of my favorite soup if it had been stirred by a used but thoroughly washed flyswatter.”  “It would bother me to sleep in a nice hotel room if I knew that a man had died of a heart attack in that room the night before.”  “If a friend offered me a piece of novelty chocolate shaped like dog-doo, I would not eat a bite.”  All of the above items rated on a scale of 1 (strongly disagree) to 9 (strongly agree) | | | | |

**Key analyses that will test the hypotheses and/or answer the research question(s).**

[Instructions: Summarise in broad ‎terms how the data will be analysed. A detailed analysis plan is not required, but the clearer the link between ‎the research question, hypotheses (as applicable), and analysis plans, the more likely the submission is to ‎pass triage. Most important are clear parameters of how to conclude this is a successful replication‎ versus to conclude this is a failed replication.]

Major analyses to be conducted:

* Two-way between-subject ANOVA tests for all DVs of experiment 1
* Three-way between-subject ANOVA tests for all DVs of experiment 2

## ‎Conclusions that will be drawn given different results.

[Instructions: Anticipate a range of possible/plausible results, what ‎they would mean for the literature, and what could be possible consequences? For example, what would it mean for the literature if this replication were to succeed, what would it mean if it were to fail?]‎

If replication succeeds: This study would provide empirical evidence for the study’s validity as well as studies that depend on this study, as well as better defining the processes through with contagion occurs

If replication fails: The above effects to the literature would not occur. Besides, alternative reasons behind why the celebrity object would need to be explored, and this may affect future research in that contagion may not be a major reason as to why certain judgement biases occur.

## Key references

[Instructions: These must be numbered and include DOI URLS. At the very least include the target for replication, and 1-2 key review/meta-analysis articles on the phenomenon from recent years.]

1. Newman et al. (2011), <https://doi.org/10.1086/658999>
2. Huang et al. (2017), <https://doi.org/10.1086/693533>

Newman et al. (2011):   
Replication and extensions

## Background

Objects that were previously owned by famous celebrities tend to fetch high prices at auction. A pair of sneakers once worn by Kanye West sold for $1.8 million US dollars in April 2021, making it the most expensive pair of sneakers that have ever been sold (Kennedy, 2021). In another example, the personal items of Ted Kaczynski (known as the Unabomber) were sold for around $190,000 USD, including a $20,053 USD price tag for a hooded sweatshirt and sunglasses among others (NBC News, 2011), showing that items which were owned by people who are widely reviled can also fetch high prices. In a third case, a tennis racket that was broken by Serena Williams during the 2018 US Open final against Osaka Naomi sold for $20,910 USD at auction (Lane, 2019). It seems clear that these objects have a high price tag attached to them not just because of their intrinsic properties, but more importantly because of their relationship to their previous owners and what those objects represented: it seems very unlikely that a mangled tennis racket would serve any practical function or purpose that would justify the over-$20,000 USD price tag. If the broken tennis racket had been swapped out with another broken racket during the auction, it seems unlikely that the buyer would go through with the purchase even if the two rackets were perfectly identical.

Newman et al. (2011) attempted to explain the reasons behind why people value objects that previously belonged to celebrities both loved and despised. Three possible reasons for this phenomenon were presented: (1) association, where celebrity objects are valuable because of who the objects are associated with; (2) market value, where celebrity objects are valuable because they might be able to be resold later on at a higher price or because people might be impressed that the owner owns these objects, or (3) the concept of contagion, the belief that a person’s “essence” can “rub off” on an object through physical contact (Newman et al, 2011).

The study described in the current report is a [very] close replication of Experiments 1 and 2 of Newman et al. (2011) and has two main goals: the first goal is to conduct independent replications of the first two experiments by replicating their methodology as closely as possible, while the second goal is to examine extensions to the study by introducing several additional measures to the experiments.

We begin by introducing the relevant literature and the chosen article for replication. Following that, we highlight the motivation for the current replication study. We then introduce several more measures to the study and assess whether these additional measures would act as mediators towards the main effects.

## Association

The first possible reason presented by Newman et al. (2011) as to why people pay money for items that were previously owned by celebrities is simply due to their sentimental association with the celebrity. As people form **positive memories** around a celebrity, Newman et al. hypothesise that they will form special memories around these celebrities that they do not want to forget. Thus, the extrinsic value of the object is that as a physical reminder, it would serve as a cue to the special memories related to that celebrity which they would want to protect (Zauberman et al., 2009), causing the object’s owner to be able to “relive [the] pleasurable emotional states” that positive memories of the celebrity would cause. However, Newman et al. (2011) also states that if this were to be the case, “objects belonging to individuals who are explicitly disliked should carry no value at all” since the association with the celebrity will then be unwanted, and the reason behind why people will pay money for their possessions will be because they “admire [these] individuals [who are generally explicitly disliked…] for whatever reason”.

## Market Value

Another possible reason may be because people think that these items are of value to others. Items which are owned by a celebrity are by definition rarer than objects that are not, which can confer to them a sense of **scarcity**. The commodity effect states that anything that can be possessed and is useful to the person possessing it will be valued according to several criteria, including scarcity (Brock, 1968; Lynn, 1992). Simply making an item more scarce or increasing its unavailability can cause a corresponding increase in its subjective value (and thus price) by others. Therefore, people may be more likely to pay higher prices for these items simply due to how scarce these items are, with the assumption that other people will pay more for these items later on; simple market forces cause these items to have a value of their own.

## Celebrity contagion

The final possible reason and the main phenomenon that Newman et al. (2011) explores is the concept of contagion, also known as magical contagion (Huang et al, 2017). Contagion in this context is a form of magical thinking in which it is believed by people that “things that once have been in contact with each other may influence each other through transfer of some of their properties via an “essence” (Rozin et al., 1986). This concept was first introduced into scientific literature by anthropologists in the late 19th century (Nemeroff and Rozin, 1994), and was used to describe various indigenous peoples’ beliefs about how things that come into physical contact with another object can transfer their properties with each other by means of a transferral of a “soul” or “mana”. This process is analogous to and is thought to have derived from a biological defense system against how germs and pathogens infect the body (Nemeroff and Rozin, 1994).

This contagion effect is shown in people who will want to avoid previously “neutral” objects which come into contact with negative sources such as a disliked person, and will be conversely attracted to items which come into contact with positive sources like a person they are sexually attracted to (Rozin et al., 1986). When applied to consumer behaviour, it has been found that in a retail setting, shoppers will rate clothing that has been previously touched by other people as being less favourable than clothing that has not been touched (Argo et al., 2006), but the reverse effect happens when the person that previously touched the clothing was an attractive salesperson of the opposite sex (Argo et al., 2008). In the case of celebrities, it has been shown that the degree of physical contact a celebrity has had with an object is positively correlated to the amount of money that people are willing to pay for it in a real-life context (Newman & Bloom, 2014).

Additionally, more recent research has also provided newer directions in which to analyse the concept of contagion. Importantly, although the review of contagion provided in Newman et al. (2011) stresses the importance of physical contact as a *conditio sine qua non* (indispensable condition)for the contagion effect to occur, more recent studies have suggested that this is not the case: for example, Smith et al. (2015) found that the serial number of an object can carry the contagion effect, as earlier serial numbers are seen to be “temporally closer to the origin (e.g., the designer or artist who produced it)” than later numbers, even when the amount of physical contact was controlled for. Huang et al. (2017) suggests that future research on the concept of contagion can focus on aspects such as the process that drives the contagion effects such as physical contact, personal valuation, and market valuation rather than aspects such as the contagion’s valence to “broaden and deepen existing models of contagion”.

## Choice of study for replication

We chose to replicate Newman et al. (2011) based on two major factors: an absence of direct replications and impact. To the best of our knowledge, there are no published direct replications of this study thus far. The article has also had a significant impact on scholarly research in the area of behavioral economics. At the time of writing, there were around 300 Google Scholar citations of the article, including many important follow-up theoretical and empirical articles with practical implications such as Newman and Bloom (2012), which takes the concepts of celebrity contagion and extends them into artwork, showing that perfect duplicates of art are less valuable than the original due to contagion; and Newmna and Dhar (2014), which extends the concept to brands and shows that “products from [an] original factory [are considered] more authentic and valuable than identical products made elsewhere”.

Should be enough for now, but keeping this section here in case more reasons are needed, since is it really necessary to first pontificate about the benefits of replication every time a replication is conducted?

(1) [What to Replicate? Justifications of study choice from 85 replication studies](https://pedermisager.netlify.com/post/what-to-replicate/), (2) Coles, N. A., Tiokhin, L., Scheel, A. M., Isager, P. M., & Lakens, D. (2018, January 17). [The Costs and Benefits of Replication Studies](https://psyarxiv.com/c8akj/) [[OSF page](https://osf.io/8gy7d/)], (3) [Quantifying the corroboration of a finding](https://pedermisager.netlify.com/post/quantifying-the-corroboration-of-a-finding/), (4) Field, S. M., Hoekstra, R., Bringmann, L., & van Ravenzwaaij, D. (2019). [When and Why to Replicate: As Easy as 1, 2, 3?](https://www.collabra.org/articles/10.1525/collabra.218/). Collabra: Psychology, 5(1)., (5) Heirene, R. (2020). [A call for replications of addiction research: Which studies should we replicate & what constitutes a “successful” replication?](https://psyarxiv.com/xzmn4/). <https://doi.org/10.31234/osf.io/xzmn4>, (6) Nosek, B. A., & Errington, T. M. (2020). [What is replication](https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000691)?. *PLOS Biology*, 18(3), e3000691. <https://doi.org/10.1371/journal.pbio.3000691>, (7) Isager, P. (2020). [What is worth replicating?](https://osf.io/ufea4/). DOI 10.17605/OSF.IO/UFEA4, 8) Isager, P. et al. (2020). [Deciding what to replicate: A formal definition of “replication value” and a decision model for replication study selection](https://osf.io/preprints/metaarxiv/2gurz/). https://doi.org/10.31222/osf.io/2gurz

We aimed to revisit this phenomenon to examine the reproducibility and replicability of the findings with independent replications. Following the recent growing recognition of the importance of reproducibility and replicability in psychological science (e.g., Brandt et al., 2014; Open Science Collaboration, 2015; van‘t Veer & Giner-Sorolla, 2016; Zwaan et al., 2018), we therefore embarked on a well-powered pre-registered [very] close replication of Newman et al. (2011).

## The original hypotheses and findings in target article

[Instructions for this section:

Hypotheses need to be directional, and you need to clearly state the expected direction. For example, writing something like "can influence" does not clearly indicate direction; instead, use words such as “increase,” “decrease,” “higher,” “lower,” “strengthen,” and “weaken”.

Since this is a replication, you also need to clearly state that you aim to replicate the original findings by XYZ study 123 and not only that there is a relationship, but in what direction the relationship would be and the expected effect-size. It needs to be clear which of the two groups you expect to be higher and to what extent.

Predictions are written in the present tense.

Make reference to the fact this is a replication, and that the effects have already received support in the original target article.

Please note: you cannot state a hypothesis that no effect or differences will be found. Unless we use Bayesian statistics, which is not in the scope of this course, there is no way for us to support the null hypothesis (no differences) over the alternative hypothesis (there are differences). This is because, when we do not find differences, we cannot be sure why. Therefore we can only reject hypotheses, not prove them. I know this is confusing, but that how Popper's NHST generally works. What we can say, is that we state the hypothesis to have differences, and then say that we do not expect to reject the null and find support for the alternative hypothesis, with an effect smaller than weak (d = 0.2) and insignificant p-values. I hope that's clear enough.]

No original hypotheses were presented by the authors in experiment 1; thus, the hypotheses for experiment 2 are presented here. Four different hypotheses were presented for experiment 2 by the original article; a summary of the hypotheses of experiment 2 is provided below in Table [X].

Table [X]

*Summary of original study hypotheses of Experiment 2*

|  |  | Positive celebrity conditions | Negative celebrity conditions |
| --- | --- | --- | --- |
| Market  value | highlighted | Purchase intentions increase (H1a) | |
| decreased | Purchase intentions decrease (H1a) | |
| Contagion | highlighted | Purchase intentions increase (H1b) | Purchase intentions decrease (H1b) |
| both  highlighted  and  decreased | Change larger in high-contagion sensitivity individuals than low-contagion sensitivity individuals (H2b) | |
| High-contagion sensitivity | | Purchase intentions increase (H2a) | Purchase intentions decrease (H2a) |

Note: associative account (pre-manipulation measures) serves as the null hypothesis for the hypotheses H1a, H1b, and H2a, but NOT H2b.

A summary of the findings in the target article is provided in Table [X]. In experiment 1, the contagion effect of valence was calculated to be = 0.39, 90% CI [0.30; 0.47], a strong effect. In experiment 2, purchase intentions was affected by the interaction of celebrity valence and physical contact ( = 0.07, 90% CI [0.03; 0.13]), a medium effect.

Table [X]

*Summary of original findings in the target article*

*Table [] Study 1 eta-squared and 90% Confidence Intervals*

|  | dfm | dfe | F | eta-squared | 90% Confidence Intervals |
| --- | --- | --- | --- | --- | --- |
| DV1: Item Value |  |  |  |  |  |
| IV: Fame | 1 | 211 | 11.42 | 0.05 | [0.01, 0.11] |
| IV: Valence | 2 | 211 | 29.48 | 0.22 | [0.14, 0.30] |
| DV2: Contagion |  |  |  |  |  |
| IV: Fame | NA | NA | NA | Unobtainable | Unobtainable |
| IV: Valence | 2 | 211 | 66.55 | 0.39 | [0.30, 0.47] |
| IV: Valence x Fame | 2 | 211 | 3.12 | 0.03 | [0.00,0.07] |

*Table [] Experiment 2 eta-squared and 90% Confidence Intervals*

|  | dfm | dfe | F | eta-squared | 90% Confidence Intervals |
| --- | --- | --- | --- | --- | --- |
| Celebrity Valence: Positive |  |  |  |  |  |
| DV1: Pleasure from Wearing  IV: Physical Contact | 1 | 111 | 12.46 | 0.10 | [0.03, 0.20] |
| DV2: Purchase Intentions  IV: Physical Contact | 1 | 111 | 17.43 | 0.14 | [0.05,0.24] |
| DV2: Purchase Intentions  IV: Market Demand | 1 | 118 | 8.35 | 0.07 | [0.01,0.15] |
| Celebrity Valence: Negative |  |  |  |  |  |
| DV1: Pleasure from Wearing  IV: Physical Contact | 1 | 108 | 12.26 | 0.10 | [0.03, 0.20] |
| DV2: Purchase Intentions  IV: Physical Contact | 1 | 108 | 2.69 | 0.02 | [0.00,0.09] |
| DV2: Purchase Intentions  IV: Market Demand | 1 | 108 | 21.24 | 0.16 | [0.07,0.28] |
| DV2: Purchase Intentions  IV: Celebrity Valence x Physical Contact | 1 | 219 | 16.77 | 0.07 | [0.03,0.13] |

## Extensions

We aimed to extend the replication of the original experiments by considering several factors that may affect the measures replicated from the original study.

Experiment 1: Alternative level of physical contact

For experiment 1, several studies conducted after the publication of the original article indicate that physical contact may not be a necessary prerequisite for contagion to occur: for example, Kim and Kim (2011) found a **proximity effect** in contagion, where an object just by being in the general vicinity of a source of “contamination” can become “infected”, without the source of “contamination” ever having to come into actual physical contact with the object itself. Stavrova et al. (2016) found that contagion can affect objects that do not even physically exist: even a piece of music can be “contaminated” by the intentions of the person who made them. Additionally, after the outbreak of the Covid-19 pandemic, participants may be averse to having physical contact with other people; as a biological view of contagion posits that contagion “originates as a defense against microbial contamination” (Nemeroff & Rozin, 1994), it may be possible that the measure of contagion in experiment 1 asking for a person’s willingness to physically come into contact with a named individual might be cross-contaminated with their perception and aversiveness towards actual microbial contamination, especially in the middle of a pandemic where the thought that “I do not want to touch other people (because of the pandemic)” may be salient. For these reasons, we added a question (“How much would you like to meet this person?”) to experiment 1 to measure the willingness of the participants to meet the person without mentioning physical contact. This extra measure will be analyzed the same as other DVs, and the results of the two-way ANOVAs are to be compared with the results of the questions measuring contagion.

Experiment 2:Perceived Market Value

For experiment 2, concern was raised that one of its main measures — specifically “the willingness to purchase the item” — may not correspond well to the measures for “market forces” in experiment 1. The measures of market value in experiment 1 (“Are there some people who would pay money for [the item]?” and “Would other people be impressed if they found out that you owned [the item]?”) was based on the participants’ perception of other people: specifically, what other people thought of the item and its value. Conversely, the “willingness to purchase the item” measure in experiment 2 was based on the participants’ self-perception: how much they themselves wanted to purchase the item and thus what its value was, instead of what other people think about the value. Due to this discrepancy in measures, the aforementioned two measures for market value in experiment 1 were copied verbatim into both the pre- and post-scenario question sets for experiment 2 to increase the validity of the measures across experiments. We combine the scores of the two questions and analyse it using a three-way ANOVA. Please refer to [Tables [X and X]](#a0cykeg4ga57) in the Methods section for more detail.

[**Note**: Apart from describing the extension with texts, it is very important to include a table for the extension design, as it enables the instructor, TAs and other peer reviewers to understand your design better. It is sometimes difficult to understand a design simply based on texts. Please check the Replication and extension experimental design table for an example.]

## Exploratory directions

[If you have a clear hypothesis it should belong to the extensions section above. If you don't, you can specify general directions here. This is also a section you can add to with things you realized after the pre-registration and the data collection. after the data has been collected with additional analyses that you haven't planned a-priori.

The pre-registration should include the regression analyses for both experiment 1 and experiment 2; these are not included here.]

## Overview of replication and extensions

The original article by Newman et al. (2011) consisted of 3 experiments; the current replication focuses on experiments 1 and 2. Participants were presented with an online survey that consisted of the two experiments presented in a random order.

For experiment 1, participants were first randomly assigned to one of six conditions in a 3 (valence: positive, negative, and mixed) by 2 (fame: celebrity and noncelebrity) between-subjects design. Each participant were first made to elicit the name of a positive, negative, or mixed valence individual (with the prompts “[someone] who you consider to be incredibly moral”, “[someone] who you consider to be incredibly immoral”, and ““[someone] who you consider to be to be of mixed moral valence; i.e., someone who is both moral at times and immoral at other times” respectively); in the celebrity conditions, the “someone” was “a living celebrity or public figure”, while in the noncelebrity conditions, the “someone” was “a living person (someone you know personally)”. After generating a name, the participant was prompted to answer a few questions related to how much they liked the person and some measures of contagion related to the person; we also extended the original experiment by asking “how much [they wanted] to meet this person”. The participant was then prompted to imagine that the individual owned one of three items (a wristwatch, a sweater, and a pair of gloves); the participant was then provided with a series of questions that measured the participant’s valuation of the object, the object’s perceived contagion, the market value of the item, and the perceived historical significance of the item. All questions were counterbalanced by presenting the questions to the participant in a random order.

In experiment 2, participants were randomly assigned to one of two conditions: one condition asked that the participant generate the name of a positive celebrity, and one where the participant was asked to generate the name of a negative celebrity. The participant was asked several questions measuring their willingness to purchase the item and the pleasantness of wearing the item. As an extension to the original experiment, the participant was also asked several questions regarding their perceived market value of the item. After providing these baseline responses, the participants were then further divided into one of four conditions: the participant was exposed to a manipulation where the contagion of the item or the market value of the item were to be manipulated by either highlighting or decreasing it, resulting in 8 conditions in a 2 (valence: positive vs. negative) by 2 (manipulation: contagion vs. market value) by 2 (direction: highlighted vs. decrease) between-subjects design. Their willingness to purchase the item, the pleasantness of wearing the item, and the perceived market value of the item were once again measured and recorded.

## Pre-registration and open-science

We first pre-registered the experiment on the Open Science Framework (OSF) and data collection was launched later that week. Pre-registrations, power analyses, and all materials used in these experiments are available in the supplementary materials. OSF pre-registration review links: [insert OSF pre-registration link here, test it to make sure that it is accessible in browser incognito mode to people who are not you].

We provided all data, JAMOVI and R/RMarkdown code for all studies on: (review link:) <https://drive.google.com/drive/folders/1NVtpRlVVb66O0iKdfyyBackYBKYgg7JC?usp=sharing>.

We provided open-science details and disclosures in the supplementary [section name, page number]. All measures, manipulations, exclusions conducted for this investigation are reported, all studies were pre-registered with power analyses reported in the supplementary [*Power analysis of original study effect to assess required sample for replication, page 22*](https://docs.google.com/document/d/1OYgFI_E-ks4khwz2hQyg4v5FKgRf0RUroDk1uyBJVm4/edit#bookmark=id.p3iukj8ij9o5), and data collection was completed before analyses.

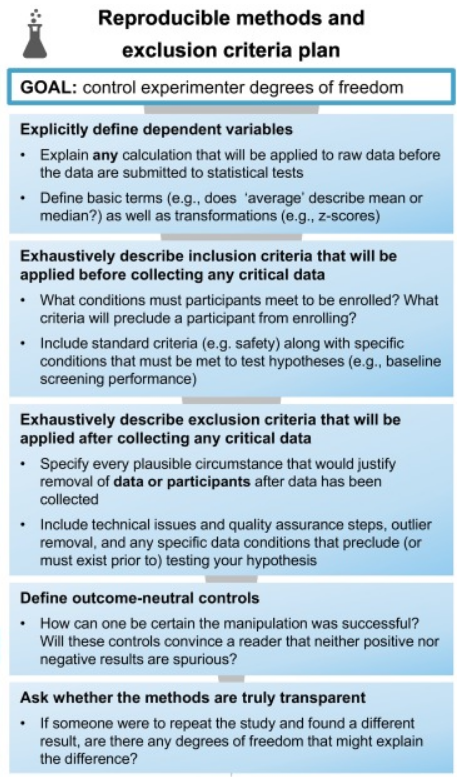
# Method

**[**Keep in Stage 1 (for pre-registration), remove after pre-registration and data collection **IMPORTANT:   
Method and results sections were written using randomized dataset produced by Qualtrics to simulate what these sections will look like after data collection. These will be updated following the data collection.]**

[Instructions for this section:

This section in the article analysis needs to be SUPER clear. The supplementary needs to be aligned with this section to include a full description of all materials, what participants saw for each one of the conditions, including, the scenario, and the questions. Everything should be labelled with proper English, no acronyms, and explained in as much detail as possible.]

[Things to keep in mind about methods from <https://www.cell.com/trends/neurosciences/fulltext/S0166-2236(19)30124-9>:]



## 

## Power analysis

To ensure that the current replication sample has sufficient power, we calculated effect sizes (ES) and power based on the statistics reported in the target article. Since the original study conducted ANOVAs, we chose Cohen’s  *f* calculated from as the effect size indicator.We set the alpha value to be 0.05 for all calculations, aiming for a power of 0.95. R and Rstudio is used to perform all calculations. We concluded that the minimum required sample size for a power of 0.95 and alpha of 0.05 is 249 participants for experiment 1 and 639 for experiment 2.‎ More detailed information regarding these calculations can be found in section [“Power analysis of original study”](https://docs.google.com/document/d/1OYgFI_E-ks4khwz2hQyg4v5FKgRf0RUroDk1uyBJVm4/edit#bookmark=id.p3iukj8ij9o5) of the supplementary.

* Experiment 1

*Table [ ]**Effect Sizes and Require Sample Size*

|  | Test Type | Eta-Squared  and 90% CI | Calculated  Cohen’s *f* | Required Sample Size |
| --- | --- | --- | --- | --- |
| DV1: Item Value |  |  |  |  |
| IV: Fame  (Celebrity vs. Non-celebrity) | Two-way between-subject ANOVA | 0.05  [0.01,0.11] | 0.23 | 249 |
| IV: Valence  (Positive vs. Negative vs. Mixed) | Two-way between-subject ANOVA | 0.22  [0.14,0.30] | 0.53 | 58 |
| DV2:Contagion |  |  |  |  |
| IV: Fame  (Celebrity vs. Non-celebrity) | Two-way between-subject ANOVA | NA | NA | NA |
| IV: Valence  (Positive vs. Negative vs. Mixed) | Two-way between-subject ANOVA | 0.39  [0.30,0.47] | 0.80 | 28 |

* Experiment 2

*Table [ ]**Effect Sizes and Require Sample Size*

|  | Test Type | Eta-Squared  and 90% CI | Calculated  Cohen’s *f* | Required Sample Size |
| --- | --- | --- | --- | --- |
| Positive Celebrity |  |  |  |  |
| DV1: Pleasure from Wearing |  |  |  |  |
| IV: Physical Contact  (Highlighted vs. Decreased) | Three-way between-subject ANOVA | 0.10  [0.03,0.20] | 0.33 | 119 |
| DV2: Purchase Intentions |  |  |  |  |
| IV: Physical Contact  (Highlighted vs. Decreased) | Three-way between-subject ANOVA | 0.14  [0.05,0.24] | 0.40 | 82 |
| IV: Market Demand  (Highlighted vs. Decreased) | Three-way between-subject ANOVA | 0.07  [0.01,0.15] | 0.27 | 175 |
| Negative Celebrity |  |  |  |  |
| DV1: Pleasure from Wearing |  |  |  |  |
| IV: Physical Contact  (Highlighted vs. Decreased) | Three-way between subject ANOVA | 0.10  [0.03,0.20] | 0.33 | 119 |
| DV2: Purchase Intention |  |  |  |  |
| IV: Physical Contact  (Highlighted vs. Decreased) | Three-way between subject ANOVA | 0.02  [0.00,0.09] | 0.14 | 639 |
| IV: Market Demand  (Highlighted vs. Decreased) | Three-way between subject ANOVA | 0.16  [0.07,0.28] | 0.44 | 70 |

## Participants

[note: some samples we collect with Americans on Amazon Mechanical Turk using [CloudResearch](https://cloudresearch.com/), some samples we recruit with British participants on [Prolific](https://prolific.co/). If you’re not sure which is here, assume US MTurk, but do not deleted the details about Prolific. If needed, we’ll change those later.]

Based on these aforementioned calculations, we recruited a total of 639 [American/British] [Amazon Mechanical Turk (MTurk) participants completed the study using TurkPrime.com/CloudResearch (Litman, Robinson, & Abberbock, 2017) / Prolific] (*Mage* = XX.XX, *SD* = XX.XX; XXX females). A comparison of the target article sample and the replication samples is provided in Table [X].

Based on these aforementioned calculations, we stimulated a sample dataset of 640 participants using the qualtrics survey build according to the original study (*Mage* = 50.70, *SD* = 30.20; 173 females). A comparison of the target article sample and the replication samples is provided in Table [X].

Table [X]

*Difference and similarities between original experiment 1 and replication*

|  | Newman et al. (2011) Experiment 1 | Stimulated Random data from Qualtrics |
| --- | --- | --- |
| Sample size | 245 | 640 | |
| Geographic origin | Not mentioned (recruited from an online database maintained by  Yale University) | Not applicable | |
| Gender | 67% female (~164 female, 81 male) | 161 males, 173 females, 306 other/did not disclose | |
| Median age (years) | Not mentioned | 52.0 | |
| Average age (years) | 35.2 | 50.7 | |
| Standard deviation age (years) | Not mentioned | 30.2 | |
| Age range (years) | Not mentioned | 100 | |
| Medium (location) | Not mentioned (presumably by computer online) | Computer (online) | |
| Compensation | Not mentioned | Not applicable | |
| Year | 2011 | 2021 | |

Table [X]

*Difference and similarities between original experiment 2 and replication*

|  | Newman et al. (2011) Experiment 2 | Stimulated Random data from Qualtrics |
| --- | --- | --- |
| Sample size | 455 | 640 | |
| Geographic origin | Not mentioned (recruited from an online database maintained by  Yale University) | Not applicable | |
| Gender | 64% female (~291 female, 164 male) | 161 males, 173 females, 306 other/did not disclose | |
| Median age (years) | Not mentioned | 52.0 | |
| Average age (years) | 34.1 | 50.7 | |
| Standard deviation age (years) | Not mentioned | 30.2 | |
| Age range (years) | Not mentioned | 100 | |
| Medium (location) | Not mentioned (presumably by computer online) | Computer (online) | |
| Compensation | Not mentioned | Not applicable | |
| Year | 2011 | 2021 | |

## 

## Design and procedure

We summarize the experimental design in Tables [X and X]. Based on our analysis of the original article, we summarize the design of experiment 1 as having a 3 (valence: positive, negative, and mixed) by 2 (fame: celebrity and noncelebrity) between-subjects design, and experiment 2 as having a 2 (valence: positive and negative) by 2 (manipulation: contagion and market value) by 2 (direction: highlighted vs. decreased) between-subjects design.

The display of experiments, conditions and specific DV items were all counterbalanced by Qualtrics’ Randomiser function to display questions in random order and the two experiments in a randomised order, and also to distribute participants randomly and evenly across the different conditions. We provide more details and all measures in the [X] section in the supplementary.

Table [X]

*Replication and extension experimental design for experiment 1*

| **IV1: Valence of named individual** [between subjects]  **IV2: Fame of named individual** [between subjects] | **IV1: Positive valence**  Participant requested to name an individual whom they consider to be “incredibly moral” | **IV1: Negative valence**  Participant requested to name an individual whom they consider to be “incredibly immoral” | **IV1: Mixed valence**  Participant requested to name an individual whom they consider to be “moral at times and immoral at other times” |
| --- | --- | --- | --- |
| **IV2: Celebrity**  Participants requested to name a celebrity or a public figure | **Dependent variables**  DV title: **Object valuation** (Primary dependent measure)  Specific DV items:  “How much would you like to own [this item]?”  “How likely would you be to purchase [this item] if it was for sale?”  “Is [this item] worth keeping, or would you throw it out?”  All of the above items rated on a scale of 1 (not at all) to 9 (very much so)  DV title: **Contagion**  Specific DV items:  “How much would you want to give this person a hug or shake their hand?”  “How much would you like to hold this item in your hands?”  All of the above items rated on a scale of 1 (not at all) to 9 (very much so)  DV title: **Market value**  Specific DV items:  “Are there some people who would pay money for this item?”  “Would other people be impressed if they found out that you owned this item?”  All of the above items rated on a scale of 1 (not at all) to 9 (very much so)  DV title: **Historical significance**  Specific DV item:  “Does this item have historical value (i.e. should it be put in a museum?)”, rated on a scale of 1 (not at all) to 9 (very much so)  DV title: **Liking**  Specific DV item:  “How much do you like this person?”, rated on a scale of 1 (extreme disliking) to 9 (extreme liking)  **Extension DV item**:  “How much would you like to meet this person?”, rated on a scale of 1 (not at all) to 9 (very much so)  Manipulation check:  “How famous is this person?”, rated on a scale of 1 (not at all famous) to 9 (extremely famous) | | |
| **IV2: Noncelebrity**  Participants requested to name a person whom they know personally |

Table [x]

*Replication and extension experimental design for experiment 2*

| **IV1: Valence of celebrity** (between subjects)  **IV2: Type of manipulation** (between subjects)  **IV3: Direction of manipulation** (between subjects) | **IV2: Manipulation of physical contact**  Participants read a description where the degree of physical contact of the celebrity with a sweater was manipulated | | **IV2: Manipulation of market demand**  Participants read a description where the market demand for the celebrity’s sweater was manipulated | |
| --- | --- | --- | --- | --- |
| **IV3: Highlighted direction**  The celebrity’s degree of physical contact with the sweater was highlighted (“it was one of their favourite sweaters”) | **IV3: Decreased direction**  The celebrity’s degree of physical contact with the sweater was decreased (“they never […] even opened the box that it came in”) | **IV3: Highlighted direction**  The sweater’s market demand was highlighted (“there is a lot of demand for items owned by [the celebrity]”) | **IV3: Decereased direction**  The sweater’s market demand was decreased (“there is very little demand for items owned by [the celebrity]”) |
| **IV1: Positive celebrity**  Participants requested to name a celebrity whom they deeply admired | **Dependent variables**  DV title: **Willingness to purchase the item**  Specific DV item:  “(Imagine that you have the opportunity to bid on a sweater that belonged to [the celebrity].) How willing are you to purchase this sweater compared to an identical used sweater (in the same condition) that was not owned by [the celebrity]?”, rated on a scale of 1 (much less likely to purchase) to 9 (much more likely to purchase)  DV title: **Pleasantness of wearing the item**  Specific DV item:  “How pleasant would you find the experience of wearing [the celebrity]'s sweater?”, rated on a scale of 1 (extremely unpleasant) to 9 (extremely pleasant)  **Extension DV items**:  “Are there some people who would pay money for this item?”  “Would other people be impressed if they found out that you owned this item?”  All of the above items rated on a scale of 1 (not at all) to 9 (very much so)  Dependent variables and extension DV items are measured once before and once after the manipulation  Manipulation checks:  “Is [the celebrity] somebody you know personally?”  “Is [the celebrity] somebody that the average person has heard of?”  All of the above items rated on a binary measure | | | |
| **IV1: Negative celebrity**  Participants requested to name a celebrity whom they deeply despised |
| **IV individual differences predictor**: **Individual contagion sensitivity**  Specific items:  “Even if I were hungry, I would not drink a bowl of my favorite soup if it had been stirred by a used but thoroughly washed flyswatter.”  “It would bother me to sleep in a nice hotel room if I knew that a man had died of a heart attack in that room the night before.”  “If a friend offered me a piece of novelty chocolate shaped like dog-doo, I would not eat a bite.” (Haidt et al., 1994)  All of the above items rated on a scale of 1 (strongly disagree) to 9 (strongly agree) | | | | |

Participants were first provided with a consent form to participate in the experiment. After their consent was obtained and verification was passed, the two experiments were then provided in a randomised order.

In experiment 1, the participants were first prompted to generate the name of an individual (the specific instructions for generating the name depending on the experimental condition they were sorted into), and were then prompted to answer a few questions related to the individual named. The participants were then prompted to imagine that the individual owned one of three objects: a sweater, a wristwatch, or a pair of gloves. The participant was then prompted to answer another set of questions related to the item that the person hypothetically owned. This entire process (the generation of a name, answering questions related to the individual, and then answering questions related to a hypothetically owned item) was then repeated two more times, such that each participant has seen and answered the same questions three times; once with a sweater, once with a wristwatch, and once with a pair of gloves. The participants were prompted to answer with another name if they had seen and completed the question set before.

In experiment 2, the participant is first prompted to provide the name of a celebrity, the valence of whom depending on the experimental condition. After going through a couple of manipulation checks, the participants are asked to imagine that they had the opportunity to bid on a sweater that belonged to the celebrity, and to answer a few questions regarding the sweater. After completing these questions, the participant is exposed to one of four different experimental manipulations, and the participant is asked to answer the same questions regarding the sweater again. The participant was then presented with several different questions that acted as predictors no matter their experimental condition.

At the end of the experiments, the participants answered a number of funneling questions, provided their demographic information, and were then debriefed. We provide a more comprehensive overview of the survey procedure in the supplementary. Further details about the predictors, manipulations, and measures used in these two experiments are provided below.

## Predictors

## Replication

For experiment 2, individual difference predictors of contagion were measured using a contagion scale adapted from Haidt et al., 1994. The scale included three items:

1. Even if I were hungry, I would not drink a bowl of my favorite soup if it had been stirred by a used but thoroughly washed flyswatter.
2. It would bother me to sleep in a nice hotel room if I knew that a man had died of a heart attack in that room the night before.
3. If a friend offered me a piece of novelty chocolate shaped like dog-doo, I would not eat a bite. (1 = *strongly disagree*, 9 = *strongly agree* for all questions; = [indicate Cronbach alpha here after data analysis]).

## Manipulations

We provide additional details of the differences in manipulation between all of the conditions, the experimental design, and the complete scales used in the current replication in the supplementary.

### Replication

#### Experiment 1, IV1:

Each participant was randomly assigned to one of three different valence conditions: a positive condition, a negative condition, and a mixed condition.

In the positive condition, participants were asked to generate the name of an individual whom they “consider to be incredibly moral”, while in the negative condition, the participants were instead asked to generate the name of an individual whom they “consider to be incredibly immoral” instead. In the mixed condition, participants were asked to generate the name of an individual whom they considered “to be of mixed moral valence; i.e., someone who is both moral at times and immoral at other times”.

#### Experiment 1, IV2:

Each participant was randomly assigned to one of two different fame conditions: a celebrity condition, and a non-celebrity condition.

In addition to the above valence manipulation when generating the name of an individual, in the celebrity condition, the name of the person generated was also requested to be “the name of a living celebrity or public figure (not someone [they knew] personally)”; in the non-celebrity condition, this was instead “the name of a living person (someone you know personally)”.

***Experiment 2, IV1:***

Each participant was randomly assigned to one of two different valence conditions: a positive condition, and a negative condition.

In the positive condition, the participant was first asked to provide the name of their “favorite living celebrity or public figure. This could be a movie star, a musician, a professional athlete, a politician, etc.” It was also specified that “this should be someone whom [they] like very much and admire and would be excited to meet personally.” In the negative condition, the participant was instead asked to “provide the name of a living person, whom [they] consider to be evil, or to personify evil; not someone [they] know personally, but a villain. This could be a mass murderer, or a fanatical leader—someone that [they] have strong negative feelings about.”

***Experiment 2, IV2/3:***

After baseline values for the DVs were obtained, each participant was then randomly assigned to one of two different types of manipulation: a manipulation of contagion, and a manipulation of market value. These two different types of manipulation were then further divided into two different directions of manipulation, a highlighting of the type of condition, or a decrease of the type of manipulation, making four different types of manipulation in total.

In the contagion+highlighted manipulation, the participant was told that a sweater owned by the celebrity “was given to [the celebrity] as a gift and it was one of their favorite sweaters and they wore it often”. In the contagion+decreased manipulation, the participant was instead told that the sweater “was given to [the celebrity] as a gift, but they never actually wore it or even opened the box that it came in”.

On the other hand, in the market value+highlighted manipulation, the participant was told that “there is a lot of demand for items owned by [the celebrity], so if [they] wanted to, it is highly likely that [they] could resell the sweater to someone else”. In the market value+decreased manipulation, the participant was told that “there is very little demand for items owned by [the celebrity], so even if [they] wanted to, it is highly unlikely that [they] could resell the sweater to someone else”. DVs were then measured again after the manipulation.

## Measures

### Replication

#### Experiment 1, Measure 1 (primary dependent measure): Object valuation

Object valuation was measured using three different items:

1. How much would you like to own [the item]?
2. How likely would you be to purchase [the item] if it was for sale?
3. Is [the item] worth keeping, or would you throw it out?

All of the above items are rated on a scale of 1 (not at all) to 9 (very much so). Scores for all three items are averaged into a single score.

#### Experiment 1, Measure 2: Contagion

Object valuation was measured using two different items:

1. How much would you want to give this person a hug or shake their hand?
2. How much would you like to hold this item in your hands?

All of the above items are rated on a scale of 1 (not at all) to 9 (very much so). Scores for both items are averaged into a single score.

#### Experiment 1, Measure 3: Market Value

Market value was measured using two different items:

1. Are there some people who would pay money for this item?
2. Would other people be impressed if they found out that you owned this item?

All of the above items rated on a scale of 1 (not at all) to 9 (very much so). Scores for both items are averaged into a single score.

#### Experiment 1, Measure 4: Historical significance

Historical significance was measured using a single item:

* Does this item have historical value (i.e. should it be put in a museum?)

Item was rated on a scale of 1 (not at all) to 9 (very much so).

#### Experiment 1, Measure 5: Liking

Liking was measured using a single item:

* How much do you like this person?

Item was rated on a scale of 1 (extreme disliking) to 9 (extreme liking).

The above DVs in Experiment 1 were measured three times per participant; the scores for each measurement were averaged to produce a single figure.

#### Experiment 2, Measure 1: Willingness to purchase item

Willingness to purchase item was measured using a single item:

* How willing are you to purchase this sweater compared to an identical used sweater (in the same condition) that was not owned by [the celebrity]?

Item was rated on a scale of 1 (much less likely to purchase) to 9 (much more likely to purchase). This DV was measured once before the manipulation and once after the manipulation.

#### Experiment 2, Measure 2: Pleasantness of wearing item

Willingness to purchase item was measured using a single item:

* How pleasant would you find the experience of wearing [the celebrity]'s sweater?

Item was rated on a scale of 1 (extremely unpleasant) to 9 (extremely pleasant). This DV was measured once before the manipulation and once after the manipulation.

### 

### Extensions

#### Experiment 1: Additional contagion measure

The extension for experiment 1, an additional measure of contagion, was measured using one item:

* How much would you like to meet this person?

Item was rated on a scale of 1 (not at all) to 9 (very much so). Similarly to the other measures in experiment 1, this item was measured three times per participant; the scores for each measurement were averaged to produce a single figure.

#### Experiment 2: Additional market value measure

The extension for experiment 2, an additional measure of market value, was measured using two different items:

* Are there some people who would pay money for this item?
* Would other people be impressed if they found out that you owned this item?

All of the above items rated on a scale of 1 (not at all) to 9 (very much so). Scores for both items are averaged into a single score. This DV was measured once before the manipulation and once after the manipulation.

## Deviations from the original article

Several minor deviations are present in the current replication. Apart from the added extensions, one of the deviations from the original text is the rewording of some questions from using the phrase “him/her” to the word “them”, the phrase “his/her” to “they”, and so on. As the individuals the participants name may be non-binary, using a more gender-neutral term would serve to be more inclusive; this is not expected to cause any significant differences in the results as a minor wording change that is predicted not to materially affect any of the DVs. This would also make it unnecessary to ask for the gender of the individuals the participants name in the study.

[Sample characteristics might be different but we won’t know until we actually know what sample we’re doing (UK or US?)]

## Evaluation criteria for replication findings

## We aimed to compare the replication effects with the original effects in experiment 1 ( = 0.39, 90% CI [0.30; 0.47]) and effect of experiment 2 ( = 0.07, 90% CI [0.03; 0.13]) using the criteria set by LeBel et al. (2019) (see section [“Replication evaluation” in the supplementary](https://docs.google.com/document/d/1OYgFI_E-ks4khwz2hQyg4v5FKgRf0RUroDk1uyBJVm4/edit#bookmark=id.j301k6t73yv4)).

## Replication closeness evaluation

We provided details on the classification of the replications using the criteria by LeBel et al., (2018) criteria in Table [X] below (see section “replication closeness evaluation” in the supplementary). We summarize the replication as a "very close” replication.

Table X

*Classification of the replication, based on LeBel et al. (2018)*

| **Design facet** | **Replication** | **Details of deviation** |
| --- | --- | --- |
| Effect/hypothesis | Same |  |
| IV construct | Similar | Very minor textual differences in the original and the replication |
| DV construct | Similar | Extra DVs were added in the form of extensions |
| IV operationalization | Same |  |
| DV operationalization | Same |  |
| Population (e.g. age) | Simulated data |  |
| IV stimuli | Same |  |
| DV stimuli | Same |  |
| Procedural details | Same |  |
| Physical settings | Presumably same | Original study also presumed to have been conducted online |
| Contextual variables | Same |  |
| Replication classification | Very close replication |  |

[Instructions for table: Read the classification in the supplementary in detail, if needed check the original article. Make a choice for each row, and only keep the right category, same, similar, or different. If similar or different, please detail what the differences are in the column “details of deviation”.]

## Data analysis strategy

### Replication: As in the original

In experiment 1, the original study averaged the ratings across the three different objects to produce one score for each type of measure (item value, contagion, market value, liking, and historical significance). For the main dependent variable (Item Value), a two-way analysis of variance (ANOVA) with alpha = 0.05 was performed to assess group differences of fame, valence and their interaction. The original study also performed hierarchical regression for analysis of mediators (contagion, market value, liking and historical value), which will not be covered in the replication.

In experiment 2, the original study calculated the difference between pre-and-post manipulation of measurables (purchase intention and pleasure from wearing). The score of changes are used as a dependent variable for three-way analysis of variance (ANOVA) with alpha = 0.05 with three IV (positive/negative celebrity, manipulation of contagion/market value, highlighted/decreased manipulation).

### Replication: Additional analyses

Considering the complexity of the original study, there is no planned additional analysis at this stage.

### Extensions

In experiment 1, in addition to the original measurement of willingness of physical contact ("​​How much would you want to give this person a hug or shake their hand?"), we added one additional measurement to provide a different degree of physical contact ("how much do you want to meet this person?"). We used this additional measurement as an extension dependent variable. The same two-way ANOVA is performed to this DV as the rest of the DVs, and results are compared to assess if the means of physical contagion makes difference to the evaluation of item value.

In experiment 2, in addition to the original measurement, we added additional measurements of the perceived market value of the item (in questions “Are there some people who would pay money for the sweater?” and “Would other people be impressed if they found out that you owned the sweater?”), drawing inspiration from the question design of experiment 1. We will calculate the change score of these two measurements and use the average score as additional DVs for the same three-way ANOVA analysis as the rest of the DVs.

### Data analyses plans

#### Outliers and exclusions

Since all responses are based on a 9-Likert point scale, it is not likely that there will be any outliers detected in any of the DVs.

#### Assumption checks

The original study did not provide information about the assumption checks nor posthoc analysis of ANOVAs. In our replication, we will implement an assumption check of homogeneity for each ANOVA, and use corresponding alternative analysis and corrections if the assumptions are violated.

# Results

**IMPORTANT:   
Method and results sections were written using randomized dataset produced by Qualtrics to simulate what these sections will look like after data collection. These will be updated following the data collection.]**

## Replication

## Experiment 1

Descriptive statistics of all measures are presented in Table [X]. Statistical tests of the hypotheses are summarized in Table [X].

Table [X]

*Descriptive statistics for all conditions - DV: Object Valuation*

|  | Valence: Positive | Valence: Negative | Valence: Mixed | Overall |
| --- | --- | --- | --- | --- |
| Fame: Celebrity | 5.08 [0.143] (107) | 4.91 [0.144] (106) | 5.00 [0.143] (107) | 5.00 [0.0827] (320) |
| Fame: Non-celebrity | 5.01 [0.144] (106) | 4.97 [0.143] (107) | 4.98 [0.143] (107) | 4.99 [0.0827] (320) |
| Overall | 5.05 [0.101] (213) | 4.94 [0.101] (213) | 4.99 [0.101] (214) | 4.99 [0.0584] (640) |

*Note*. [The data in the table represent Mean [SE] (No. of participants) respectively.]

Figure [x]. Descriptive statistics and pair-wise comparison of DV1: Item Value

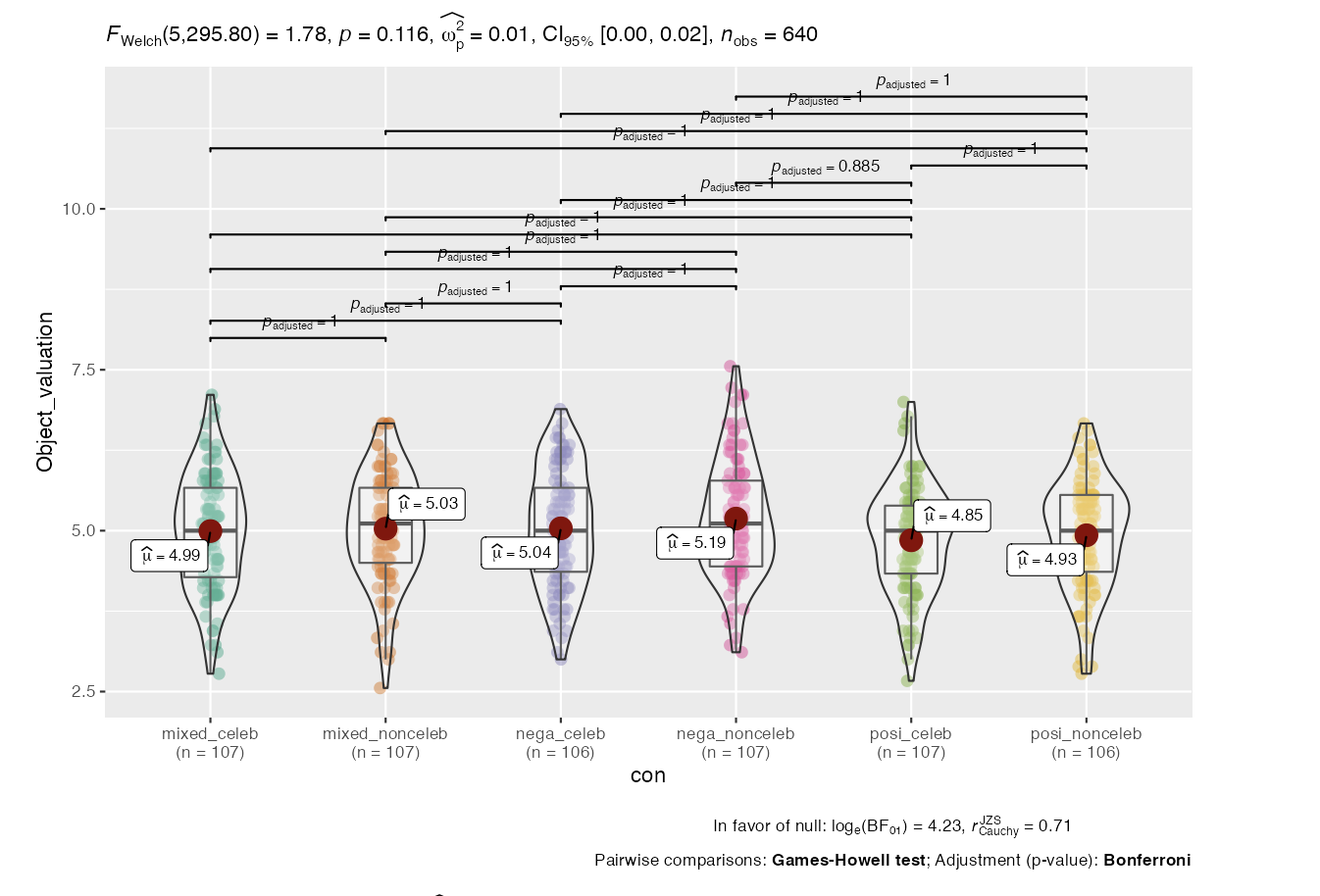


Table [X] *Summary of statistical tests*

|  | | dfm | | dfe | F-stat | p | η² | Interpretation |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Fame |  | 1 |  | 638 | 1.642 | 0.201 | 0.003 | [no-signal] [inconsistent]  [smaller than original] |
| Valence |  | 2 |  | 637 | 3.437 | 0.033 | 0.011 | [no-signal] [inconsistent]  [smaller than original] |
| Fame ✻ Valence |  | 2 |  | 634 | 0.230 | 0.794 | 0.001 | [no-signal] [inconsistent]  [smaller than original] |
| *Note*. Please refer to the [cloud folder](https://drive.google.com/drive/folders/1NVtpRlVVb66O0iKdfyyBackYBKYgg7JC?usp=sharing) for the complete statistics and visualization for other DVs. | | | | | | | | |

We conducted two-way ANOVA on object valuation with valence (Positive: *N* = 213, *M* = 5.05, *SD* = 0.101; Negative: N = 213, *M* = 4.94, *SE* = 0.101; Mixed: N = 214, *M* = 4.99, *SE* = 0.101) and fame (Celebrity: *N* = 320; *M* = 5.00, *SE* = 0.0827; Non celebrity: *M* = 4.99, *SE* = 0.0827) as dependent variables. Though there is small effect of valence on object valuation (*p* = .033;  *= 0.003*) nor t, there is no effect of fame (*p* = .201;  *= 0.011*). We failed to find support for the hypothesis that fame and ethical valence affect the item value mediated by contagion.

In comparison, the original study found the contagion effect of valence was = 0.39, 90% CI [0.30; 0.47], a strong effect, which our simulated data did not replicate.

## Experiment 2

Descriptive statistics of all measures are presented in Table [X]. Statistical tests of the hypotheses are summarized in Table [X].

Table [X]

*Descriptive statistics for positive celebrity - DV: Change of purchase intention*

|  | Contagion | Market Value |
| --- | --- | --- |
| Highlighted | -0.550 [0.423] (80) | -0.125 [0.423] (80) |
| Decreased | -0.313 [0.423] (80) | 0.313 [0.423] (80) |

*Note*. [The data in the table represent Mean [SE] (No. of participants) respectively.]

*Descriptive statistics for Negative celebrity - DV: Change of purchase intention*

|  | Contagion | Market Value |
| --- | --- | --- |
| Highlighted | 0.800 [0.405] (80) | 0.438 [0.405] (80) |
| Decreased | -0.162 [0.405] (80) | 0.637 [0.405] (80) |

*Note*. [The data in the table represent Mean [SE] (No. of participants) respectively.]

Table [X]

| *Summary of three-way ANOVA on change of purchase intention* | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | dfm | | dfe | F | | p | η² | Interpretation |
| valence | 1 |  | 638 | 4.155 |  | 0.042 | 0.006 | [no-signal] [inconsistent]  [smaller than original] |
| manipulation | 1 |  | 638 | 1.613 |  | 0.205 | 0.003 | [no-signal] [inconsistent]  [smaller than original] |
| direction | 1 |  | 638 | 0.006 |  | 0.940 | 0.000 | [no-signal] [inconsistent]  [smaller than original] |
| valence ✻ manipulation | 1 |  | 636 | 0.273 |  | 0.601 | 0.000 | [no-signal] [inconsistent]  [smaller than original] |
| valence ✻ direction | 1 |  | 636 | 1.506 |  | 0.220 | 0.002 | [no-signal] [inconsistent]  [smaller than original] |
| manipulation ✻ direction | 1 |  | 636 | 1.353 |  | 0.245 | 0.002 | [no-signal] [inconsistent]  [smaller than original] |
| valence ✻ manipulation ✻ direction | 1 |  | 632 | 0.675 |  | 0.412 | 0.001 | [no-signal] [inconsistent]  [smaller than original] |
|  | | | | | | | | |

*Note*. Please refer to the [cloud folder](https://drive.google.com/drive/folders/1NVtpRlVVb66O0iKdfyyBackYBKYgg7JC?usp=sharing) for the complete statistics and visualization.

We conducted a three-way ANOVA on purchase intention using celebrity valence, manipulation of contagion/market value, and manipulation direction as IVs. Though there is small effect of valence (*p* = .042;  *= 0.006*), we failed to find support for the hypothesis that manipulation of contagion and market value would affect evaluation of items owned by celebrities.

In comparison, the original study found that purchase intentions was affected by the interaction of celebrity valence and physical contact ( = 0.07, 90% CI [0.03; 0.13]), a medium effect.

## Extensions

* Experiment 1

We added an extension DV to measure the willingness of an alternative degree of physical contact. The descriptive statistics are summarized in Table [X].

Table [X]

*Descriptive statistics for all conditions - DV: Extension of Contagion*

|  | Valence: Positive | Valence: Negative | Valence: Mixed | Overall |
| --- | --- | --- | --- | --- |
| Fame: Celebrity | 5.01 [0.142] (107) | 5.02 [0.142] (106) | 5.02 [0.142] (107) | 5.01 [0.0822] (320) |
| Fame: Non-celebrity | 4.72 [0.143] (106) | 5.06 [0.142] (107) | 5.13 [0.142] (107) | 4.97 [0.0822] (320) |
| Overall | 4.86 [0.101] (213) | 5.04 [0.101] (213) | 5.07 [0.101] (214) | 4.99 [0.0584] (640) |

*Note*. [The data in the table represent Mean [SE] (No. of participants) respectively.]

The result of the two-way ANOVA of extension DV is summarized in the table [X].

| Table [X] | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
|  |  | dfe | dfm |  | F |  | p | η² |
| fame |  | 1 | 638 |  | 0.140 |  | 0.709 | 0.000 |
| valence |  | 2 | 637 |  | 1.239 |  | 0.290 | 0.004 |
| fame ✻ valence |  | 2 | 634 |  | 1.097 |  | 0.335 | 0.003 |
|  | | | | | | | | |

* Experiment 2

We added an extension DV of market value, same as experiment 1. The results of the three-way ANOVA are summarized in Table [X].

| Table [X] |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | df |  | F |  | p |  | η² |
| Overall model |  | 7 |  | 0.2709 |  | 0.965 |  |  |
| valence |  | 1 |  | 0.0443 |  | 0.833 |  | 0.000 |
| manipulation |  | 1 |  | 0.1649 |  | 0.685 |  | 0.000 |
| direction |  | 1 |  | 0.3440 |  | 0.558 |  | 0.001 |
| valence ✻ manipulation |  | 1 |  | 0.9557 |  | 0.329 |  | 0.002 |
| valence ✻ direction |  | 1 |  | 0.0111 |  | 0.916 |  | 0.000 |
| manipulation ✻ direction |  | 1 |  | 0.3619 |  | 0.548 |  | 0.001 |
| valence ✻ manipulation ✻ direction |  | 1 |  | 0.0145 |  | 0.904 |  | 0.000 |
| Residuals |  | 632 |  |  |  |  |  |  |
|  | | | | | | | | |

## Comparing replication to original findings

For the simulated data, we failed to replicate the original findings; the responses were very similar across conditions, as expected from the computer generalized random dataset.

# Discussion

[Please note that the discussion is only to be completed in Stage 2 following data collection]

We conducted a pre-registered replication of [Phenomenon name]. The results are [consistent/not consistent/partially consistent and partially inconsistent] with the original results (see Table X for a summary of the replication).

## Replication

Overall, [if applicable, mention the strengths of this replication, e.g. with a larger well-powered sample and diverse sample], we found that: (1) [...] , (2) [...] , (3) [...], and (4) [...]

In summary, goal of the project was to assess the replicability of the research presented by [Target article author1] & [Target article author2] ([target article year]) in support of [Phenomenon name].

[Link back to the literature, and perhaps discuss the theoretical implications of the findings. If applicable, discuss the practical implications of the findings, but do not infer too much]

## Extensions

We ran extensions examining [...]. Overall, our findings showed that [...] We found [weak to no / weak / medium / strong] support for our hypothesis. [Discussion of the extension findings, linking to the literature]

## Implications, limitations, and directions for future research

### Constraints on Generality (COG)

[Please read [Constraints on Generality (COG): A Proposed Addition to All Empirical Papers](https://pdfs.semanticscholar.org/56b4/b04e13911edf1bfec00b45febef168425181.pdf) (Simons et al., 2017) for more details about this paragraph and examples.

A COG statement should take these principles into account when specifying the target populations and constraints for each of the following aspects of their study.

Participants. Discuss how your proximal sample of participants is representative of a broader target population. If you tested undergraduates, do you believe the findings apply only to students at your university? To students at comparable universities? To students anywhere? To all adults? To all mammals? What would another researcher need to do to verify that their participants were drawn from the same target population as yours? If your COG statement specifies a target population of all adults, you are accepting that a replication with any sample of adults would constitute a direct test of the same effect.

Materials/Stimuli. Define the class of materials/stimuli to which your finding should generalize. What are the critical features of your materials that must be maintained to measure the same construct? What measurements are necessary to verify that any new materials tap the same target population as your materials?

Procedures. What aspects of your procedures must be followed closely to reproduce the effect? What broader class of procedures should produce the same results? For example, would future studies need to test participants in your lab? would researchers need to use the same computers, and if so, have you provided enough detail in your method section for them to do so? Will the effect work only if participants are tested in isolated cubicles, or will it also work in a large classroom setting or in a shopping mall? Can any undergraduate administer the tasks or does the study require special training? What checks are needed to ensure that the procedures match the broader population of procedures that can produce the effect?

Historical/temporal specificity. Does the effect depend on cultural norms that might change over time? For example, results of studies involving attitudes about same-sex marriage in a study conducted in the 1990s might differ from those in a study conducted in 2017. Similarly, studies of attitudes about politics might differ when measured right before or after an election. What aspects of the temporal or historical context need to be stable to observe the effect? Can you anticipate and specify any differences in the historical or temporal context that might affect whether or not other researchers would observe the same effect?]

[Example1](https://journals-sagepub-com.eproxy.lib.hku.hk/doi/10.1177/1745691617708630#): “Our finding provides evidence of the Dunning-Kruger effect (Kruger & Dunning, 1999) in participants who are aware of their relative skill. Given that this “better than average” effect has been observed for a diverse range of participants in a wide range of tasks (including unpublished evidence from our own laboratory with chess players), we expect our result with bridge players to generalize to other domains in which players regularly compete against the same group of players in games of skill. However, given that relative performance in any given session of duplicate bridge involves some luck, the pattern of results—optimistic predictions but accurate memory—might hold only for games that involve both skill and luck. A direct replication would test bridge players in sessions that include players with skill levels ranging from relative novice to expert in the context of their regular bridge game (i.e., the players should play with and against each other at least weekly and should be familiar with the skill level of the other players in each session). Participants should be blind to the predictions made by other players to avoid having knowledge of those predictions affect their play. We have no reason to believe that the results depend on other characteristics of the participants, materials, or context. “

[Example 2](https://journals-sagepub-com.eproxy.lib.hku.hk/doi/10.1177/1745691617708630#): “The stimuli consisted of a large number of video clips in which a large number of different undergraduates sampled from the subject pool at the University of Washington each expressed mild distress in their own way. Thus, we expect the results to generalize to situations in which participants view similar video clips, as long as manipulation checks indicate the clips depict a variety of ways in which people express mild distress. Unpublished studies from our laboratory resulted in similar results despite variations in the testing context (e.g., different research assistants). Consequently, we do not expect such variations to matter. We believe the results will be reproducible with students from similar subject pools serving as participants. However, we do not have evidence that the findings will occur outside of laboratory settings. The distress expressed in the video clips was triggered by a specific laboratory induction, and we lack evidence showing that the results will generalize to expressions of distress in response to other situations. We have no reason to believe that the results depend on other characteristics of the participants, materials, or context.”

[Example 3](https://journals-sagepub-com.eproxy.lib.hku.hk/doi/10.1177/1745691617708630#): “The results from our no-photo condition converge with prior evidence that combining a plausible narrative attributed to a family member with social pressure, demand characteristics, and sustained memory recovery techniques can lead a substantial percentage of undergraduate subjects to appear to remember a childhood pseudoevent. The relative contributions of these components is unclear. Moreover, the likelihood of false memory reports is affected by numerous variables including the nature of the suggested event (see Lindsay & Read, 2001); the absolute rate of false memories in our study should not be used to predict the probability of false memories of childhood sexual abuse. Moreover, the very high false memory rate in our photo condition may be specific to this suggested event and photo. Our suggested event involved an accomplice, and we speculate that this may have amplified the photo effect by helping subjects imagine the event. We do expect, though, that the rates of false memory for similar types of events (i.e., events with a similar rate of false memory) should generally be higher with a photo memory prompt than without one provided that the photo supplies information that participants can use to imagine the suggested event. It must be noted, however, that our sample sizes were modest, especially given the nature of the measures and the design, so the absolute rates of false memories that we observed might well differ in replications on statistical grounds. We speculate that asking subjects about increasingly remote events (a Grade 5 or 6 event and then a Grade 3 or 4 event before asking about the Grade 1 or 2 pseudoevent) may also have increased false memory rates. Finally, all subjects were tested by the second author, who was (in the judgment of the first author) adept at presenting the suggestions in a compelling way and motivating the subjects (who were younger than she) to work hard at remembering the pseudoevent. We speculate that these skills increase the likelihood of false memory reports. We have no reason to believe that the results depend on other characteristics of the participants, materials, or context.”

### Constraints on theory generalizability

[Note: this is following the debate on crisis of generalization, see Twitter thread here: <https://twitter.com/lakens/status/1259130713510739968> .

See also:  
Yarkoni, T. (2019, November 22). [The Generalizability Crisis](https://psyarxiv.com/jqw35). <https://doi.org/10.31234/osf.io/jqw35>  
Fried, E. I. (2020, February 7). [Lack of theory building and testing impedes progress in the factor and network literature](https://psyarxiv.com/zg84s/). https://doi.org/10.31234/osf.io/zg84s]

We observed X, finding [support/mixed support/no support] for the findings in the original article. [If you found mixed support, describe on which DV(s)/scenario(s)/hypothesis/hypotheses you found support and on which DV(s)/scenario(s)/hypothesis/hypotheses you failed to find support. This may be related to boundary conditions of the phenomenon.] Given the link suggested between the theory/phenomenon and these experimental designs and findings, we interpret this to be in [support/mixed support] of the theory/phenomenon in this specific context and methodology. We note that our ability to generalize from these findings to other contexts, such as [examples of other contexts], and methodology, such as [examples of other methods], is limited and implications for theory needs to be further elaborated and tested.

### Constraints on population generalizability

[Note: this is following the debate on constraint about generalizing from WEIRD samples to overall population world wide, see Twitter thread here: <https://twitter.com/harrisonmanley/status/1272031885623885826?s=20>.

See paper: Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). [Beyond western, educated, industrial, rich, and democratic (WEIRD) psychology: measuring and mapping scales of cultural and psychological distance](http://eprints.lse.ac.uk/104663/1/Beyond_WEIRD_Psychology_Main.pdf). *Psychological Science*, 0956797620916782. ]

We conducted the study with a sample from X [...]. We note that our ability to generalize from this sample to other samples is limited, and generalizability to other samples needs to be further elaborated and tested. [...You may mention that the generalizability to non-WEIRD samples may be uncertain or unclear, if applicable, with reference to studies/articles related to this phenomenon, theories/concepts explain this phenomenon. You may also call for multi-lab and multi-country collaborations/replications with more diverse samples, e.g. Psychological Science Accelerator (Moshontz et al., 2018) and Many labs (Klein et al., 2018)]

# Conclusion

[Summarize things - 1) Main takeaways, 2) Replication findings, whether the findings are consistent with original findings, 3) Extension findings, 4) One sentence about most important future research directions]

# References

[Add references. Verify citations with [reciteworks.com](https://reciteworks.com/) , add DOIs with [crossref](https://doi.crossref.org/simpleTextQuery), and check retractions with [scite](https://scite.ai/#reference-check). Example below.   
Notes: APA style format, journals' name capitalized, must include DOI (reciteworks and crossreg help with that), if preprint/resource, add and verify link.]

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science*, *1*, 389-402.

LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A brief guide to evaluate replications. *Meta-Psychology*, 3, 1-9. DOI: 10.15626/MP.2018.843

Heirene, R. (2020, January 7). A call for replications of addiction research: Which studies should we replicate & what constitutes a “successful” replication?. <https://doi.org/10.31234/osf.io/xzmn4>

Nosek, B. A., & Errington, T. M. (2020). What is replication?. *PLOS Biology*, 18(3), e3000691. <https://doi.org/10.1371/journal.pbio.3000691>

See if you want to keep any of the following:

Isager, P. M. (2019). [Quantifying Replication Value: A formula-based approach to study selection in replication research](https://www.psycharchives.org/handle/20.500.12034/2024). *Open Science 2019, Trier, Germany*.

Field, S. M., Hoekstra, R., Bringmann, L., & van Ravenzwaaij, D. (2019). [When and Why to Replicate: As Easy as 1, 2, 3?](https://www.collabra.org/articles/10.1525/collabra.218/). *Collabra: Psychology, 5*(1).

Coles, N. A., Tiokhin, L., Scheel, A. M., Isager, P. M., & Lakens, D. (2018). [The costs and benefits of replication studies](https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/costs-and-benefits-of-replication-studies/65E3C8C7DAF24F06FE58E800A4F4DA69). *Behavioral and Brain Sciences*, *41*.

Isager, P. M. (2019). [What to Replicate? Justifications of study choice from 85 replication studies](https://pedermisager.netlify.com/post/what-to-replicate/).

Isager, P. M., van Aert, R. C. M., Bahník, Š., Brandt, M. J., DeSoto, K. A., Giner-Sorolla, R., … Lakens, D. (2020, September 2). Deciding what to replicate: A formal definition of “replication value” and a decision model for replication study selection. https://doi.org/10.31222/osf.io/2gurz

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams Jr, R. B., Alper, S., ... & Batra, R. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science, 1*(4), 443-490. <https://journals.sagepub.com/doi/full/10.1177/2515245918810225>

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime. com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior research methods*, 49(2), 433-442.

Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., ... & Castille, C. M. (2018). The Psychological Science Accelerator: Advancing psychology through a distributed collaborative network. *Advances in Methods and Practices in Psychological Science, 1*(4), 501-515. <https://doi.org/10.1177/2515245918797607>

Olsson-Collentine, A., van Assen, M. A. L. M., & Wicherts, J. M. (2020). Heterogeneity in direct replications in psychology and its association with effect size. *Psychological Bulletin*. <https://doi.org/10.1037/bul0000294>.

---Added references below, sorry for the mess---

Newman, G. E., Diesendruck, G., & Bloom, P. (2011). Celebrity contagion and the value of objects. *Journal of Consumer Research*, *38*(2), 215-228.

Kennedy, M. (2021, April 26). *Kanye West sneakers fetch record $1.8m at Private Sale*. AP News. Retrieved October 29, 2021, from <https://apnews.com/article/entertainment-music-lifestyle-business-hip-hop-and-rap-f4e3c76ce47c5f6a1806597daaac1e6e>

Lane, B. (2019, December 10). *The racket which Serena Williams smashed during her 2018 US Open final defeat to Naomi Osaka just sold at auction for $20,910.* Insider. Retrieved October 29, 2021, from <https://www.insider.com/serena-williams-smashed-racket-us-open-sells-for-20910-2019-12>.

Huang, J. Y., Ackerman, J. M., & Newman, G. E. (2017). Catching (up with) magical contagion: A review of contagion effects in consumer contexts. *Journal of the Association for Consumer Research, 2*(4), 430-443.

Brock, T. C. (1968). Implications of commodity theory for value change. In A. G. Greenwald, T. C. Brock, & T.M. Ostrom (Eds.), Psychological foundations of attitudes (pp. 243-275). New York: Academic.

Lynn, M. (1992). The Psychology of Unavailability: Explaining Scarcity and Cost Effects on Value. *Basic and Applied Social Psychology, 13*(1), 3–7. <https://doi.org/10.1207/s15324834basp1301_2>

NBC News. (2011, June 3). *Unabomber auction nets $190,000*. <https://www.nbcnews.com/id/wbna43257976>

Zauberman, G., Ratner, R. K., & Kim, B. K. (2009). Memories as Assets: Strategic Memory Protection in Choice over Time. *Journal of Consumer Research*, *35*(5), 715–728. https://doi.org/10.1086/592943

Kim, L. R., & Kim, N. S. (2011). A proximity effect in adults' contamination intuitions. *Judgment and Decision making*, *6*(3), 222.

Nemeroff, C., & Rozin, P. (1994). The Contagion Concept in Adult Thinking in the United States: Transmission of Germs and of Interpersonal Influence. *Ethos*, *22*(2), 158–186. https://doi.org/10.1525/eth.1994.22.2.02a00020

Smith, R. K., Newman, G. E., & Dhar, R. (2015). Closer to the Creator: Temporal Contagion Explains the Preference for Earlier Serial Numbers. *Journal of Consumer Research, 42(5), 653–668. https://doi.org/10.1093/jcr/ucv054*

Savani, K., Kumar, S., Naidu, N. V. R., & Dweck, C. S. (2011). Beliefs about emotional residue: the idea that emotions leave a trace in the physical environment. *Journal of Personality and Social Psychology, 101*(4), 684–701. <https://doi.org/10.1037/a0024102>

Rozin, P., Millman, L., & Nemeroff, C. (1986). Operation of the laws of sympathetic magic in disgust and other domains. *Journal of Personality and Social Psychology, 50*(4), 703–712. https://doi.org/10.1037/0022-3514.50.4.703

Argo, J. J., Dahl, D. W., & Morales, A. C. (2006). Consumer Contamination: How Consumers React to Products Touched by Others. *Journal of Marketing*, *70*(2), 81–94. https://doi.org/10.1509/jmkg.70.2.081

Argo, J. J., Dahl, D. W., & Morales, A. C. (2008). Positive Consumer Contagion: Responses to Attractive Others in a Retail Context. *Journal of Marketing Research, 45*(6), 690–701. https://doi.org/10.1509/jmkr.45.6.690

Newman, G. E., & Bloom, P. (2014). Physical contact influences how much people pay at celebrity auctions. *Proceedings of the National Academy of Sciences*, *111*(10), 3705–3708. https://doi.org/10.1073/pnas.1313637111

Haidt, J., McCauley, C., & Rozin, P. (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual differences*, *16*(5), 701-713.

Stavrova, O., Newman, G. E., Kulemann, A., & Fetchenhauer, D. (2016). Contamination without contact: An examination of intention-based contagion. *Judgment and Decision making, 11*(6), 554.