[Link to supplementary: Purple [Group B] - Replication & Extension - Supplementary]

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

Mannix Chan
Department of Psychology, University of Hong Kong, Hong Kong SAR
mannix42@connect.hku.hk / mannixchan9149@gmail.com

Yaqi Jin
Department of Psychology, University of Hong Kong, Hong Kong SAR
jinyaqi@connect.hku.hk / yvonne27hku@gmail.com

*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.

Devin CHEUNG and Kristy Chow were teaching assistants at the University of Hong Kong psychology department during the academic year 2021-2022.

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 interest with respect to the authorship and/or publication of this article.

Financial disclosure/funding:

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

Authorship declaration:

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

Devin CHEUNG guided and assisted the replication effort in the PSYC2071.

Gilad Feldman was the course instructor for PSYC2071 and led the replication efforts. He supervised each step in the project.

Corresponding author

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

Full details of all authors:

First name	Last name	Researchgate profile	ORCID	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	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@connect.hku.hk	yvonne27hku@gmail.com	3035535339

Contributor Roles Taxonomy

In the table below we 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) for details and definitions of each of the roles listed below.

Table [X]

Role	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	
Funding acquisition	X			
Investigation	X			X
Pre-registration peer review / verification		X	X	X
Data analysis peer review / verification				X
Methodology	X	X	X	
Project administration	X			X
Resources	X			
Software	X	X	X	
Supervision	X			X
Validation	X			
Visualization	X		X	
Writing-original draft	X	X	X	
Writing-review and editing	X	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	https://osf.io/m3s9e/?view_only=9ec49acb1fe74f02b9db39536c5018cb
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

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.

We thank Dr. Aaron Charlton for reviewing and providing insightful comments on an early version of this manuscript, and for Dr. George Newman for providing us with materials used in the original experiments and for reviewing the Qualtrics survey. <- pending

This is a revised version after receiving immensely helpful feedback from external reviewers, course TAs and peer reviews from teammates. For a recorded list of what we have changed and improved from the initial submission, please refer to this notion page: https://cut-cantaloupe-d72.notion.site/Replication-manuscript-Final-revised-submission-beccee2bd25d4649a6384a429ca98a75

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 , add DOIs with crossref.
3. Run statcheck.
4. See RRR assessment template 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.]

Items that were previously owned by famous celebrities can fetch high prices. Several reasons behind these high prices may include memory association, market value, or contagion (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 (info here), 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

The main hypothesis of experiment 1 is that 1) intention to engage in physical contact increases for items belonging to positive celebrities and non-celebrity compared to those belonging to negative celebrities and 2) contagion and market forces are significant predictors of the valuation of celebrity items. Specifically, positive celebrity possessions value increases when either market value or contagion increases, whereas the value of possessions of negative celebrity increases when market value increases and decreases when contagion increases.

The original/replication hypothesis for experiment 2 can be found in the table below.

<table>
<tr><td></td><td></td><td>Positive celebrity valence</td><td>Negative celebrity valence</td></tr>
<tr><td rowspan="2">Market demand</td><td>highlighted</td><td colspan="2">Purchase intentions increase (H1a)</td></tr>
<tr><td>decreased</td><td colspan="2">Purchase intentions decrease (H1a)</td></tr>
<tr><td rowspan="2">Physical contact</td><td>highlighted</td><td>Purchase intentions increase (H1b)</td><td>Purchase intentions decrease (H1b)</td></tr>
<tr><td>both highlighted and decreased</td><td colspan="2">Change larger in high-contagion sensitivity individuals than low-contagion sensitivity individuals (H2b)</td></tr>
<tr><td colspan="2">High-contagion sensitivity</td><td>Purchase intentions increase (H2a)</td><td>Purchase intentions decrease (H2a)</td></tr>
</table>

Study design and methods

Both experiments were conducted with an online survey.

Experiment 1:

<table>
<tr><td>IV1: Valence of named individual [between subjects]
IV2: Fame of named individual [between subjects]</td><td><u>IV1: Positive valence</u>
Participant requested to name an individual whom they consider to be “incredibly moral”</td><td><u>IV1: Negative valence</u>
Participant requested to name an individual whom they consider to be “incredibly immoral”</td><td><u>IV1: Mixed valence</u>
Participant requested to name an individual whom they consider to be “moral at times and immoral at other times”</td></tr>
<tr><td><u>IV2: Celebrity</u>
Participant requested to name a celebrity or a public figure</td><td colspan="3"><u>Dependent variables</u>
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?”
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)</td></tr>
<tr><td><u>IV2: Noncelebrity</u>
Participant requested to name a person whom they know personally</td></tr>
</table>

	<u>Extension DV item</u>: "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)

Experiment 2:

<table>
<tr><td rowspan="2"></td><td colspan="2"><u>IV2: Manipulation of physical contact</u>
Participants read a description where the degree of physical contact of the celebrity with a sweater was manipulated</td><td colspan="2"><u>IV2: Manipulation of market demand</u>
Participants read a description where the market demand for the celebrity's sweater was manipulated</td></tr>
<tr><td><u>IV3: Highlighted direction</u>
The celebrity's degree of physical contact with the sweater was highlighted ("it was one of their favourite sweaters")</td><td><u>IV3: Decreased direction</u>
The celebrity's degree of physical contact with the sweater was decreased ("they never […] even opened the box that it came in")</td><td><u>IV3: Highlighted direction</u>
The sweater's market demand was highlighted ("there is a lot of demand for items owned by [the celebrity]")</td><td><u>IV3: Decereased direction</u>
The sweater's market demand was decreased ("there is very little demand for items owned by [the celebrity]")</td></tr>
<tr><td><u>IV1: Positive celebrity</u>
Participant requested to name a celebrity whom they deeply admired</td><td colspan="4" rowspan="2"><u>Dependent variables</u>
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)

<u>Extension DV items</u>:</td></tr>
<tr><td><u>IV1: Negative celebrity</u>
Participant requested to name a celebrity whom they deeply despised</td></tr>
</table>

	"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
<u>IV individual differences predictor</u>: **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).

The main analyses include:

- 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.

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 on 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

Theoretical 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). Items that were owned by people who are widely reviled can also fetch high prices. The personal items of Ted Kaczynski (known as the Unabomber) were sold for around $190,000 USD (NBC News, 2011). Even a broken tennis racket used by Serena Williams during the 2018 US Open final against Osaka Naomi was 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.

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 original article found the last one to be of the most explanatory power.

The study described in the current report is a 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 to explore additional factors regarding perceptions on contagion and market value.

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 clue to the special memories related to that celebrity which they would want to protect (Zauberman et al., 2009). However, Newman et al. (2011) also state 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 that people think that these items are of value to others. Items that 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) explore 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 defence 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 that 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. 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. They cite the example of the original vinyl pressing for the Beatles' *White Album*; approximately three million were produced, making its scarcity relatively low, but vinyls with earlier serial numbers sold for higher prices than did those with later ones. The contagion effect is observed even though the artists themselves may not have come into the vinyl pressings themselves, since they are ostensibly factory-produced objects. Huang et al. (2017) suggest 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: the absence of direct replications and their impact. To the best of our knowledge, there are no published direct replications of this study thus far. At the time of writing, there were around 300 Google Scholar citations of the article, and the article has also had a significant impact on scholarly research in the area of behavioural economics, 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". The success of this replication would also have practical implications in furthering the understanding of the market forces

behind the trade of items connected to celebrities. Due to these reasons, we therefore embarked on a well-powered pre-registered close replication of Newman et al. (2011).

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, (2) Coles, N. A., Tiokhin, L., Scheel, A. M., Isager, P. M., & Lakens, D. (2018, January 17). The Costs and Benefits of Replication Studies [OSF page], (3) 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?. 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://doi.org/10.31234/osf.io/xzmn4, (6) Nosek, B. A., & Errington, T. M. (2020). What is replication?. *PLOS Biology*, 18(3), e3000691. https://doi.org/10.1371/journal.pbio.3000691, (7) Isager, P. (2020). What is worth replicating?. 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://doi.org/10.31222/osf.io/2gurz

The original hypotheses and findings in the target article

No original hypotheses were explicitly 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 1.

Table 1

Summary of original study hypotheses of Experiment 2

		Positive celebrity valence	Negative celebrity valence
Market demand	highlighted	Purchase intentions increase (H1a)	
	decreased	Purchase intentions decrease (H1a)	
Physical contact	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, including those effects studied in experiment 1, is provided in Tables 2 and 3. In experiment 1, the contagion effect of valence was calculated to be $\eta^2 = 0.39$, 90% CI [0.30; 0.47], a strong effect. In experiment 2, purchase intentions were affected by the interaction of celebrity valence and physical contact ($\eta^2 = 0.07$, 90% CI [0.03; 0.13]), a medium effect.

Table 2

Experiment 1 ANOVA analysis summary

	dfm	dfe	*F*	Partial eta-squared	90% Confidence Intervals
DV1: Object Valuation					
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 3

Experiment 2 ANOVA analysis summary

	dfm	dfe	F	Partial 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	1	108	2.69	0.02	[0.00,0.09]

IV: Physical Contact					
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]

Note: for the comprehensive summary and analysis of the original article, please refer to the Original article results section in the supplementary.

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. Two extensions were added to the study; one measured contagion without physical contact, and one served as both a measure of perceived market value and as a manipulation check.

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 defence 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. Besides, in some cases, physical contagion may not occur with people where handshaking and/or hugging is not culturally prevalent for people who are not familiar with each other. For example. those from some Islamic traditions believe that physical contact with the opposite sex should be discouraged or should be prohibited (IslamFYI, 2017), and may thus not display an effect of physical contagion towards another person when framed as "shaking hands with the person" or "hugging the person" due to this effect being mediated by personal religious and/or cultural beliefs. 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.

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, specifically in the highlighted/decreased contagion conditions. 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. This causes a disparity in measures between the two experiments for the contagion conditions. On the other hand, in the highlighted/decreased market value conditions, since the perceived market value of the item is manipulated, adding a measure of perceived market value would serve as a manipulation check to see whether the manipulations were successful. Given these reasons, 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 in the highlighted/decreased contagion conditions. Please refer to Tables 7 and 8 in the Methods section for more details.

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 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 was 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, their willingness to come into contact with the item, 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 random order.

In experiment 2, participants were randomly assigned to one of two conditions: one condition asked that the participant generate the name of their favourite living celebrity or public person, and one where the participant was asked to generate the name of a famous living person they believed to be evil. 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 celebrity's physical contact with the item (contagion) and the market value of the item were both manipulated. This resulted in 8 conditions in a 2 (celebrity valence: positive vs. negative) × 2 (contagion: highlighted vs. decreased) × 2 (market value: highlighted vs. decreased) between-subjects design. Participants' willingness to purchase the item, the pleasantness of wearing the item, and the perceived market value of the item were once again measured.

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 link: https://osf.io/m3s9e/?view_only=9ec49acb1fe74f02b9db39536c5018cb

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. 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*, 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.]

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 partial η^2 as the effect size indicator. We set the alpha value to be 0.05 for all calculations, aiming for a power of 0.95. 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. R code used and the summary tables of these calculations can be found in the Effect size calculations of the original study effects section and Power analysis of the original study of the supplementary.

Table 4

Effect Sizes and Required Sample Size for Experiment 1

	Test Type	Partial Eta-Squared and 90% CI	Calculated Cohen's *f*	Required Sample Size
DV1: Object Valuation				
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	Two-way	0.39	0.80	28

(Positive vs. Negative vs. Mixed)	between-subject ANOVA	[0.30,0.47]

Table 5

Effect Sizes and Required Sample Size for Experiment 2

	Test Type	Partial 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	0.16 [0.07,0.28]	0.44	70

ANOVA

Participants

Based on these aforementioned calculations, we stimulated a sample dataset of 640 participants using the Qualtrics survey build according to the original study (M_{age} = 50.70, *SD* = 30.20; 173 females). A comparison of the target article sample and the replication samples is provided in Table 6.

Table 6

Comparison of participant pools between original experiment 1, experiment 2 and replication

	Newman et al. (2011) Experiment 1	Newman et al. (2011) Experiment 2	Simulated random data from Qualtrics
Sample size	245	455	640
Geographic origin	Not mentioned, recruited from an online database maintained by Yale University		Not applicable
Gender	67% female (164 female, 81 male)	64% female (291 female, 164 male)	161 males, 173 females, 306 other/didn't disclose
Median age (years)	Not mentioned		52.0
Average age (years)	35.2	34.1	50.7
SD of age (years)	Not mentioned		30.2
Range of age (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 7 and 8. Based on our analysis of the original article, we summarize the design of experiment 1 as having a 3 (celebrity valence:

positive, negative, or mixed) × 2 (fame: celebrity or noncelebrity) between-subjects design, and experiment 2 as having a 2 (celebrity valence: positive vs. negative) × 2 (contagion: highlighted vs. decreased) × 2 (market value: highlighted vs. decreased) between-subjects design.

The display of experiments, conditions and specific DV items were all counterbalanced by Qualtrics' Randomiser function to display certain questions in random order and to display the two experiments in a randomised order, and also to distribute participants randomly and evenly across the different conditions. We provide an overview of the mechanics of the Qualtrics survey in a README file in the cloud folder.

Table 7

Replication and extension experimental design for experiment 1

<table>
<tr><td>IV1: Valence of named individual [between subjects]
IV2: Fame of named individual [between subjects]</td><td><u>IV1: Positive valence</u>
Participant requested to name an individual whom they consider to be “incredibly moral”</td><td><u>IV1: Negative valence</u>
Participant requested to name an individual whom they consider to be “incredibly immoral”</td><td><u>IV1: Mixed valence</u>
Participant requested to name an individual whom they consider to be “moral at times and immoral at other times”</td></tr>
<tr><td><u>IV2: Celebrity</u>
Participants requested to name a celebrity or a public figure</td><td colspan="3"><u>Dependent variables</u>
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?”
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)

<u>Extension DV item</u>:
“How much would you like to meet this person?”, rated on a scale of 1 (not at all) to 9 (very much so)</td></tr>
<tr><td><u>IV2: Noncelebrity</u>
Participants requested to name a person whom they know personally</td></tr>
</table>

	Manipulation check: "How famous is this person?", rated on a scale of 1 (not at all famous) to 9 (extremely famous)

Table 8

Replication and extension experimental design for experiment 2

<table>
<tr><td rowspan="2">IV1: Valence of celebrity (between subjects)
IV2: Type of manipulation (between subjects)
IV3: Direction of manipulation (between subjects)</td><td colspan="2"><u>IV2: Manipulation of physical contact</u>
Participants read a description where the degree of physical contact of the celebrity with a sweater was manipulated</td><td colspan="2"><u>IV2: Manipulation of market demand</u>
Participants read a description where the market demand for the celebrity's sweater was manipulated</td></tr>
<tr><td><u>IV3: Highlighted direction</u>
The celebrity's degree of physical contact with the sweater was highlighted ("it was one of their favourite sweaters")</td><td><u>IV3: Decreased direction</u>
The celebrity's degree of physical contact with the sweater was decreased ("they never […] even opened the box that it came in")</td><td><u>IV3: Highlighted direction</u>
The sweater's market demand was highlighted ("there is a lot of demand for items owned by [the celebrity]")</td><td><u>IV3: Decereased direction</u>
The sweater's market demand was decreased ("there is very little demand for items owned by [the celebrity]")</td></tr>
<tr><td><u>IV1: Positive celebrity</u>
Participants requested to name a celebrity whom they deeply admired</td><td colspan="4" rowspan="2"><u>Dependent variables</u>
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).

<u>Extension DV items</u>:
"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.</td></tr>
<tr><td><u>IV1: Negative celebrity</u>
Participants requested to name a celebrity whom they deeply despised</td></tr>
</table>

	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.
<u>IV individual differences predictor</u>: **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) <u>All of the above items rated on a scale of 1 (strongly disagree) to 9 (strongly agree)</u>	

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 presented in random order.

In experiment 1, 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), then prompted to answer a few questions related to the individual they named. Participants were then prompted to imagine that the individual owned one of three randomly-chosen objects: a sweater, a wristwatch, or a pair of gloves. The participant then answered 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 answered the same questions three times; once with a sweater, once with a wristwatch, and once with a pair of gloves. Participants were prompted to answer with a

different name than they had used previously if it was not their first time seeing the set of questions throughout the experiment.

In experiment 2, participants were 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 were asked to imagine that they had the opportunity to bid on a sweater that belonged to the celebrity, and were prompted to answer a few questions regarding the sweater. After completing these questions, the participant was exposed to one of four different experimental manipulations, and the participant was asked to answer the same questions regarding the sweater again. All participants were then presented with a 3-item sensitivity to contagion individual differences scale (modified from Haidt et al., 1994).

At the end of the experiments, the participants answered a number of funneling questions, including how serious they were in filling out the questionnaire and what they thought the purpose of the study was. Participants provided their demographic information, and were then debriefed.

Manipulations

Experiment 1, IV1 (valence):

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 (fame):

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 (celebrity valence):

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 (type of manipulation/direction of manipulation):

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

For experiment 1, object valuation was measured using three different questions, namely "How much would you like to own the item?", "How likely would you be to purchase [the item] if it was for sale?", and "Is the item worth keeping?" All of the above items were rated on a Likert scale of 1 (not at all) to 9 (very much so). Scores for all three items were averaged into a single score. Contagion was measured using two different items: "How much would you want to give this person a hug or shake their hand?", and "How much would you like to hold this item in your hands?". Both of the above items were rated on a scale of 1 (not at all) to 9 (very much so). Scores for both items are averaged into a single score. Market value was measured using two different items: "Are there some people who would pay money for this item?" and "Would other people be impressed if they found out that you owned this item?". Both of the above items were rated on a scale of 1 (not at all) to 9 (very much so). Scores for both items are averaged into a single score. 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). 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.

For experiment 2, the willingness to purchase the 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]?". This item was rated on a scale of 1 (much less likely to purchase) to 9 (much more likely to purchase). The pleasantness of wearing the item was

measured using a single item: “How pleasant would you find the experience of wearing [the celebrity]'s sweater?”. This item was rated on a scale of 1 (extremely unpleasant) to 9 (extremely pleasant). Both of the above DVs were measured once before the manipulation and once after the manipulation.

The extension for experiment 1, an additional measure of contagion, was measured using one the question “How much would you like to meet this person?”. This 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.

The extension for experiment 2, an additional measure of market value and a manipulation check, was measured using two different items: “Are there some people who would pay money for this item?” and “Would other people be impressed if they found out that you owned this item?”. Both of these items were 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 modernisation of some questions by shifting from the use of gendered pronouns: "him/her," "his/hers" to gender-neutral pronouns: "them," "they." As the individuals the participants' name may be non-binary, using a more gender-neutral term is more inclusive. This wording change is not expected to cause material differences in any of the results. This change would also make it unnecessary to ask for the gender of the individuals the participants name in the study.

Another change made to the study is a modification to one of the questions in experiment 1 regarding object valuation. One of the questions that measured object valuation in Newman et al. (2011) was "Is this item worth keeping, or would you throw it out?", measured on a Likert scale of 1 (not at all) to 9 (very much so). The meaning of this question is unclear upon first glance as it seems like a binary question (either keep it or throw it out), and it becomes uncertain whether the answer "1 (not at all)" would mean "this item is not at all worth keeping" or "I do not want to throw this item out at all". To make this question clearer to the participants, this question has been modified to read "Is this item worth keeping?". The method of analysis remains unchanged for this question in this replication.

[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 **(η^2**= 0.39, 90% CI [0.30; 0.47]**)** and effect of experiment 2 (η^2 = 0.07, 90% CI [0.03; 0.13]) using the criteria set by LeBel and colleagues (2019) (see **Replication evaluation in the supplementary**)**.**

Replication closeness evaluation

We provided details on the classification of the replications using the criteria by LeBel et al., (2018) criteria in Table 9 below (see section “replication closeness evaluation” in the supplementary). We classify the replication as a "very close” replication.

Table 9

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; see section on deviations from the original article
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 (object valuation, contagion, market value, liking, and historical significance). The main analysis includes a two-way analysis of variance (ANOVA) on the main dependent variable (object valuation) to the effect of fame, valence and their interaction. Further analysis on other measures (contagion, market value, liking and historical significance) is also reported in an incomplete fashion. 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 on all DVs with IVs being fame and valence, and results are compared to assess if the means of physical contagion makes a difference in item valuation.

In experiment 2, in addition to the original measurements, 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 post-hoc 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 10. Statistical tests of the hypotheses are summarized in Table 11.

Table 10

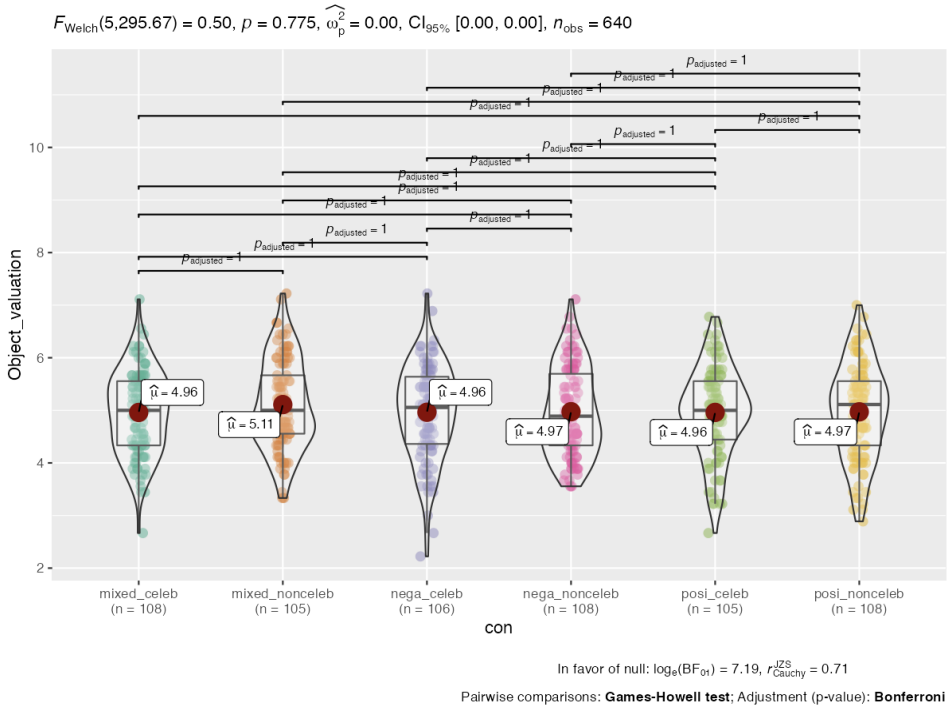
Descriptive statistics for all conditions - DV: Object Valuation

	Valence: Positive	Valence: Negative	Valence: Mixed	Overall

Fame: Celebrity	4.96 [0.844] (105)	4.96 [0.891] (106)	4.96 [0.143] (108)	4.96 [0.0840] (319)
Fame: Non-celebrity	4.97 [0.897] (108)	4.97 [0.827] (108)	5.11 [0.848] (105)	5.02 [0.0857] (321)
Overall	5.05 [0.101] (213)	4.94 [0.101] (213)	4.99 [0.101] (214)	4.99 [0.0584] (640)

Note. Notation formula of the number is Mean [SD] (No. of participants).

Figure [x]. Descriptive statistics and pair-wise comparison of DV1: *Object Valuation*



Note. Please refer to the additional results section in the supplementary for the pair-wise analysis and visualization for other DVs.

Table 11 *Summary of statistical tests - DV: Object Valuation*

	dfm	dfe	*F*-stat	*p*	η^2	Interpretation

Fame	1	638	0.714	0.398	0.001	[no-signal] [inconsistent] [smaller than original]
Valence	2	637	0.442	0.643	0.001	[no-signal] [inconsistent] [smaller than original]
Fame ✻ Valence	2	634	0.471	0.625	0.001	[no-signal] [inconsistent] [smaller than original]

Note. Please refer to the JAMOVI file "Exp1_two_way_ANOVAs_for_all_DVs.omv" in the cloud folder for the ANOVA results for other DVs.

The central finding of the original article in Experiment 1 was the interaction between valence and fame on contagion. However, we failed to find support for the hypothesis that fame and ethical valence affect the object valuation mediated by contagion.

The two-way ANOVA on object value with valence and fame and their interaction as predictors does not yield a significant result, $F(2, 634) = 0.471$, $p=.625$, $\eta^2 = 0.001$. In comparison, the original study found the effect of valence on willingness to touch was $\eta^2 = 0.39$, 90% CI [0.30; 0.47], a strong effect, which our simulated data did not replicate.

Replication: Experiment 2

Descriptive statistics of all measures are presented in Table 12. Statistical tests of the hypotheses are summarized in Table 13.

Table 12

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] (sample size) 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] (sample size) respectively.]

Table 13

Summary of three-way ANOVA on change of purchase intention

	dfm	dfe	F	*p*	η^2	Interpretation
valence	1	638	0.110	0.741	0.000	[no-signal] [inconsistent] [smaller than original]
manipulation	1	638	1.210	0.272	0.002	[no-signal] [inconsistent] [smaller than original]
direction	1	638	0.795	0.373	0.001	[no-signal] [inconsistent] [smaller than original]
valence ✻ manipulation	1	636	2.593	0.108	0.004	[no-signal] [inconsistent] [smaller than original]
valence ✻ direction	1	636	0.196	0.658	0.000	[no-signal] [inconsistent] [smaller than original]
manipulation ✻ direction	1	636	0.977	0.323	0.002	[no-signal] [inconsistent] [smaller than original]
valence ✻ manipulation ✻ direction	1	632	0.182	0.670	0.000	[no-signal] [inconsistent] [smaller than original]

Note. Please refer to the "Exp2_three_way_ANOVAs_for_all_DVs.omv" in the cloud folder for the complete statistics and visualization of all other DVs.

We conducted a three-way ANOVA on purchase intention using celebrity valence, manipulation of contagion/market value, and manipulation direction as IVs. We failed to find support for the hypothesis that manipulation of contagion and market value would affect the evaluation of items owned by celebrities, and the analysis yields nearly no effect across the board. In comparison, the original study found that purchase intentions were affected by the interaction of celebrity valence and physical contact ($\eta^2 = 0.07$, 90% CI [0.03; 0.13]), a medium effect.

Extension: 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 14

Table 14 *Descriptive statistics for all conditions - DV: Extension of Contagion*

	Valence: Positive	Valence: Negative	Valence: Mixed	Overall
Fame: Celebrity	4.79 [1.66] (105)	5.27 [1.47] (106)	5.12 [1.49] (108)	5.06 [1.55] (319)
Fame: Non-celebrity	5.12 [1.45] (108)	5.15 [1.43] (108)	5.04 [1.56] (105)	5.10 [1.47] (321)
Overall	4.95 [1.56] (213)	5.21 [1.45] (214)	5.08 [1.52] (213)	5.08 [1.52] (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 table 15.

Table 15

Two-way ANOVA result for DV: Extension of Contagion

	dfe	dfm	*F*	*p*	η^2
fame	1	638	0.111	0.739	0.000
valence	2	637	1.546	0.214	0.005
fame ✻ valence	2	634	1.489	0.226	0.005

The two-way ANOVA analysis yields non-significant results across the board and there is no interaction between valence and fame, $F(2,634) = 1.489$, $p = 0.226$, $\eta^2 = 0.005$.

Extension: Experiment 2

We added an extension DV of market value, with measurements consistent with experiment 1. The results of the three-way ANOVA are summarized in Table 16.

Table 16

Two-way ANOVA result for DV: Extension - Market Value

	df	F	p	η^2
Overall model	7	1.4596	0.179	
valence	1	0.000	0.982	0.000

manipulation	1	0.268	0.605	0.000
direction	1	3.630	0.057	0.006
valence ✻ manipulation	1	3.310	0.069	0.005
valence ✻ direction	1	0.046	0.830	0.000
manipulation ✻ direction	1	0.019	0.891	0.000
valence ✻ manipulation ✻ direction	1	3.456	0.064	0.005
Residuals	632			

Similarly, in this simulated data, the three-way ANOVA analysis yields non significant results across the board and there is no interaction between valence, manipulation and direction, $F(3,632) = 3.456$, $p = 0.982$, $\eta^2 = 0.005$.

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 (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: “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: “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: “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://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://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. *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 , add DOIs with crossref, and check retractions with scite. 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.]

See if you want to keep any of the following:

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

Isager, P. M. (2019). Quantifying Replication Value: A formula-based approach to study selection in replication research. *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?. *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. *Behavioral and Brain Sciences*, *41*.

Isager, P. M. (2019). What to Replicate? Justifications of study choice from 85 replication studies.

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

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

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

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

These are the references for the intro and methods section (sorry for the messy order):

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

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. https://doi.org/10.15626/MP.2018.843

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.

IslamFYI. (2017, September 13). Why Do Some Muslims Refrain from Shaking Hands with People of the opposite gender? | IslamFYI: an Educational Resource on Islam for the Public. Islamfyi.princeton.edu. https://islamfyi.princeton.edu/why-do-some-muslims-refrain-from-shaking-hands-with-people-of-the-opposite-gender/