

RESEARCH ARTICLE

Avatars, Disconnecting Agents: Exploring the Nuances of the Avatar Effect in Online Discourse

Chris Macdonald^{1*}¹Skylab, United Kingdom*Corresponding author: Chris Macdonald: chrismacdonald@dr.com

Citation: Macdonald C. (2020) Avatars, Disconnecting Agents: Exploring the Nuances of the Avatar Effect in Online Discourse. Open Science Journal 5(2)

Received: 23rd February 2020

Accepted: 6th April 2020

Published: 8th May 2020

Copyright: © 2020 This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The author received no specific funding for this work

Competing Interests: The author has declared that no competing interests exist.

Abstract:

A recurring theme in the study of unethical and harmful behavior is disconnection. And therefore perhaps it is no surprise that when individuals or groups view themselves as a separate entity, they can loosen or even lose their moral compass. This phenomenon is known as the avatar effect. The experiment reported here explores the avatar effect in online communities. It explores whether participants adopt the suggested traits of a given entity when they communicate via an online persona. The results revealed that participants' youtube comments were significantly influenced by their username and profile image, despite being asked to give their own opinion. The paper offers further insight into the growing field of disconnecting agents and compromised ethics.

Keywords: Ethics, Psychology, Economics, Cognition, Sociology, Connection, Communication, Disconnecting agents, Avatar effect, Youtube, Internet, Trolling, Commenting, Anonymity, Pseudonymity

Introduction

A recurring theme in the study of unethical and harmful behavior is disconnection (Latané et al, 1969; Valentine, 1980; Bandura, 1992; Baillon et al, 2012; Brody et al, 2016; Macdonald, 2019; Macdonald, 2020a; Macdonald, 2020b). And therefore perhaps it is no surprise that when individuals or groups view themselves as a separate entity, they can loosen or even lose their moral compass. This phenomenon is known as the avatar effect (Macdonald, 2020c).

When making decisions through a seemingly separate entity there is the potential for anonymity and pseudonymity. The potential negative influence of anonymity is somewhat apparent. If one is not being observed then one could act

with a reduced fear of repercussions or the judgement of others (Harrison, 2015). However, anonymity doesn't necessarily reduce another key aspect that can also prevent unethical conduct: judgement from one's self.

As noted by Albert Bandura, a key component of ethical conduct is the self-regulatory mechanisms tied to personal standards (Bandura, 2001). Therefore, while we might try to avoid acting in a way that others may criticize, we also try to avoid self-condemnation (Bandura, 2002). Accordingly, people generally strive to maintain a positive self-image, even when unobserved by others (Allport, 1955; Rosenberg, 1979; Shalvi et al, 2015). Thus being able to conceal one's identity from others might not necessarily increase unethical behaviour as one may still fear the judgement from themselves.

Pseudonymity, however, could break this final barrier, as potentially, one might be able to disassociate themselves from the outcome of their own actions, and perhaps accordingly their own self-judgement.

Therefore, if one acts through an avatar (a new persona), and that avatar is not seen as a representation of themselves, but rather a separate entity altogether, with its own ideas, goals, and purpose, then we might have very fertile ground for unethical conduct as one might experience a reduction of the sense of both judgement from others and one's self.

When studying the literature on cyberbullying, there are suggestions of a possible avatar effect. It has been observed that online attackers feel less of a sense that they are doing wrong (Gini, 2006; Perren and Gutzwiller-Helfenfinger, 2012; Menesini et al, 2013). They have also been reported to feel less guilt and compassion (Menesini and Camodeca, 2008; Pozzoli and Gini, 2010). Perhaps they are able to disassociate themselves from their own tendencies and their own self-judgement. Perhaps in addition to feeling less observed, they also feel less 'themselves'.

I wondered if when communicating via online platforms, where we create usernames and profile images, could there be an inherent component where we might be unconsciously adopting a separate identity. Especially when we consider that often people are forced to use a name different from their birth name due to availability. And people often stick with default profile images for convenience. Could there be an element of pseudonymity in addition to anonymity, could there be an avatar effect? Could this help to explain scenarios where people act in a way that is uncharacteristically unethical?

The YouTube experiment

This experiment explores the avatar effect in online communities. In particular, it explores whether participants adopt the suggested traits of a given entity when they communicate via that entity. In other words, can profile choices influence whether we communicate as ourselves or 'another entity'?

Participants

96 US Students took part in the experiment. In an attempt to reduce variables, all participants also had the following in common: aged 18-25, Caucasian ethnicity, and English-only spoken at home. All participants were paid

for their participation in the experiment and none of the participants were allowed to partake in the experiment more than once.

Procedure

The students were randomly assigned to one of three groups: Group P, Group NA, and Group PA. There were 96 participants in total, 32 participants per group.

Before starting the task, each participant was reminded that there is no right or wrong answer and that they should give their honest opinion.

The participants then read through a hypothetical scenario via a private online portal. They were told to imagine that they are logged into a Youtube account and that they have just finished watching a video of someone singing and it was apparent that the singer was not very good. They were then asked to give an example of a comment that they would leave.

The scenario for each group was the same, the only difference was the details of the Youtube account:

- For Group P (Personal), the displayed username is the participant's personal name and the profile image is a picture of their face.
- For Group NA (Neutral Avatar), the displayed username is YoutubeAccount-99 and the profile image is the default image (a blank face).
- For Group PA (Primed Avatar), the displayed username is SuperMeanCritic-99 and the profile image is a cartoon image of a mean face.

After the experiment, each comment was ranked as either negative, mild, or positive. If the participant stated that they wouldn't leave a comment, or it was unclear, it was ranked as mild. Mild comments also included neutral statements, such as those that were both positive and negative, such as "The music was very good but I did not like the quality of the video".

*A note on the experiment delivery method.

A known concern with priming experiments is that the researchers conducting the experiment might be able to influence the results (Pfungst, 1965; Rosenthal, 1966; Rosenthal et al, 1969; Rosenthal et al, 1975; Barber, 1978; Conty et al, 2010; Salkind, 2010; Dreisbach et al, 2011; Doyen et al, 2012; Bargh et al, 2014; Gilder et al, 2018) even without intending to do so. To protect against this, a third-party UK organization was used to distribute the experiments. The third-party uses an automated distribution service and they were unaware of the purpose of the experiment or any hypotheses attached to it. The results were then sent back via a secure online portal.

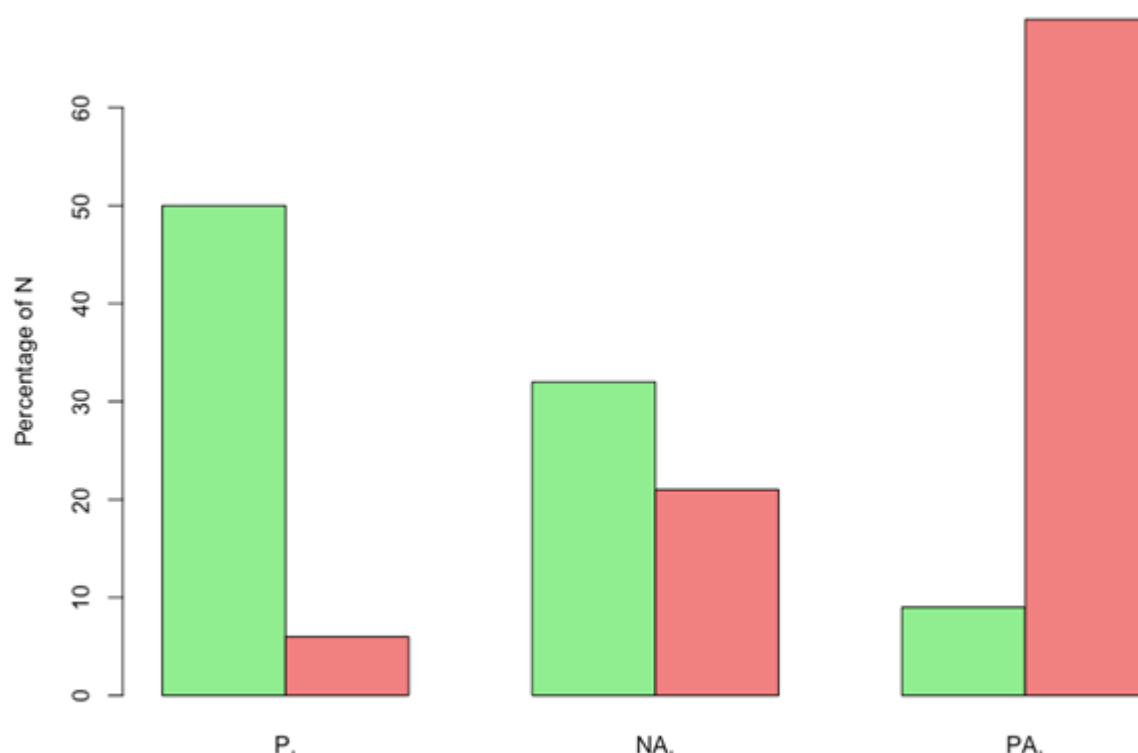
Results

There was a highly significant difference between Group P (N = 32) and Group PA (N = 32), $t(61) = -6.42$, $p < .001$.

Group P: 50% Positive 06% Negative

Group PA: 09% Positive 69% Negative

The bar graph below illustrates a clear trend whereby the percentage of negative comments increases as we move from Group P to Group NA to Group PA. Accordingly, there is also a trend whereby the percentage of positive comments decreases as we move from Group P to Group NA to Group PA.



* Green = Positive, Red = Negative

The data indicates that the participants who were given a personalized username and profile image (Group P) were significantly more positive than participants who were given the negatively primed profile image and username (Group PA). Participants with a neutral username and blank profile image were in between. That is to say that although every participant was asked to give their own comment, the username and profile image appears to have a significant influence.

Below are some examples of positive and negative comments from each group.

Personalized (personal name as username)

Group P had 16 positive comments. A recurring theme was support and encouragement; examples include, “Excellent effort!”, “Nice work, just a bit more

practice”, “It's awesome that you're putting yourself out there, keep working to improve and you'll be awesome one day!” and “Love your passion, keep at it friend”. Group P only had two negative comments: “Yikes” and “That video was horrible”. Both negative comments are short, quite general, and not personal.

Neutral Avatar (YoutubeAccount-99 as username)

Group NA had ten positive comments, and they also tended to offer encouragement and support; examples include: “Good work! Getting better every day”, “You are really making progress”, and “Love the confidence! Keep on practicing!”. Group NA had seven negative comments; examples include, “wow trash”, “this sucks”, and “I don't like the way you sing”. The negative comments tended to be slightly harsher than those given by Group P.

Primed Avatar (SuperMeanCritic-99 as username)

Group PA had only three positive comments, “Keep practicing and you'll get better every day! Best of luck”, “I give you kudos for trying”, and “You look nice”. Group PA had 22 negative comments. Not only was the negativity far more frequent, but it also tended to be significantly more personal and harsher; examples include, “stop trying to sing”, “you f***** suck”, “My disappointment is immeasurable”, “ur ugly”, and “you should be ashamed of yourself”.

Discussion

A common stance when discussing cyberbullying, trolling, and other forms of online negative discourse, is that the attackers are facilitated by anonymity (Suler, 2004; Ybarra et al, 2004; Vanderbosch et al, 2008; Kowalski et al, 2013; Peebles, 2014; Bartlett et al, 2016; Janopaul-Naylor et al, 2019). That is to say that they may be negative people in their everyday lives but as they are identifiable they maintain reasonably respectable conduct. One argument, therefore, is that when they login to an online account, they could become anonymous, and therefore they might show their 'true colours'.

However, the experiment reported here suggests that there may be something more sophisticated going on. While anonymity did seem to result in greater negativity, a primed avatar appeared to have a greater impact:

- Group P vs Group NA: $t(60) = -2.02, p < .05$
- Group P vs Group PA: $t(61) = -6.42, p < .001$

Therefore if there are situations where individuals can become a new 'entity', we could see a larger increase in unethical behavior than if the individual merely views themselves as an anonymous version of themselves.

Future research.

While this paper has focused on the negative side of the avatar effect, it would be interesting to explore the positive side of the avatar effect. Would participants become more ethical when they 'become' a positively primed entity?

I would also like to see a great deal more research into the avatar effect with corporate practices as this is a known area where individuals can act as separate entities and is also an area notorious for unethical and harmful decision-making (Lincoln et al, 1982; Murphy et al, 1992; Gunthorpe, 1997; Bowie, 1999; Koneswaran et al, 2008; Yuhao, 2010; Kennedy et al, 2013; Wiid et al, 2013; Soltani, 2014; DeAngelis, 2016; Sudsakorn et al, 2018).

Concluding remarks

When reading through online comments one could be forgiven for losing faith in humanity. Online platforms seemingly reveal the ugliest sides of us *Homo sapiens*. I believe the results reported here help to somewhat restore faith. The results suggest that people are actually rather ethical and pleasant when they communicate as themselves. They appear to have a tendency to offer support and encouragement. When individuals can communicate anonymously then negativity appears to increase, although it is still predominantly positive. The most negative comments seemed to appear when participants were most removed from themselves. Further supporting the growing evidence that disconnection may be at the heart of unethical conduct (Macdonald, 2019; Macdonald, 2020a; Macdonald, 2020b). Therefore, if we want more authentic and positive experiences, perhaps we are better off bypassing disconnecting agents.

References:

- Allport GW. 1955. *Becoming: Basic considerations for a psychology of personality*. New Haven, CT: Yale University Press
- Baillon A, Asli S, Van Dolder D. 2012. On the Social Nature of Eyes: The Effect of Social Cues in Interaction and Individual Choice Tasks, *Evolution and Human Behavior* 34:(2) 146-154 ssrn.com/abstract=2205072
- Bandura A. 1992. Social cognitive theory of social referencing, in: S. FEINMAN (Ed.) *Social Referencing and the Social Construction of Reality in Infancy*. New York: Plenum 175-208
- Bandura A. 2001. Social cognitive theory: an agentic perspective, *Annual Review of Psychology* 52: 1-26 doi.org/10.1146/annurev.psych.52.1.1
- Bandura A. 2002. Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education* 31:(2) 101-119 https://web.stanford.edu/~kcarmel/CC_BehavChange_Course/readings/Additional%20Resources/Bandura/bandura_moraldisengagement.pdf
- Barber TX. 1978. Expecting expectancy effects: Biased data analyses and failure to exclude alternative interpretations in experimenter expectancy research. *Behavioral and Brain Sciences* 1:(3) 388-390 doi.org/10.1017/s0140525x00075531
- Bargh JA, Chartrand TL. 2014. *The mind in the middle: A practical guide to priming and automaticity research*. In H. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd ed). New York: Cambridge University Press
- Barlett CP, Gentile DA, Chew C. 2016. Predicting cyberbullying from anonymity. *Psychology of Popular Media Culture* 5: 171-180 doi.org/10.1037/ppm0000055
- Bowie NE. 1999. *Business ethics: A Kantian perspective*. Malden, MA: Blackwell
- Brody N, Vangelisti AL. 2016. Bystander intervention in cyberbullying. *Communication Monographs* 83:(1) 94-119 doi.org/10.1080/03637751.2015.1044256
- Conty L, Gimmig D, Belletier C, George N, Huguet P. 2010. The cost of being watched: Stroop interference increases under concomitant eye contact. *Cognition* 115: 133-139 doi.org/10.1016/j.cognition.2009.12.005
- DeAngelis CD. 2016. Big Pharma Profits and the Public Loses. *The Milbank quarterly* 94:(1) 30-33 doi.org/10.1111/1468-0009.12171
- Doyen S, Klein O, Pichon CL, Cleeremans A. 2012. Behavioral Priming: It's All in the Mind, but Whose Mind? *PLoS ONE* 7:(1) e29081 doi.org/10.1371/journal.pone.0029081

- Dreisbach G, Boettcher S. 2011. How the social-evaluative context modulates processes of cognitive control. *Psychological Research* 75:(2) 143-151 doi.org/10.1007/s00426-010-0298-z
- Gilder TSE, Heerey EA. 2018. The Role of Experimenter Belief in Social Priming. *Psychological Science* 29:(3) 403-417 doi.org/10.1177/0956797617737128
- Gini G. 2006. Social cognition and moral cognition in bullying: What's wrong? *Aggressive Behaviour* 32 528-539 doi.org/10.1002/ab.20153
- Gunthorpe DL. 1997. Business Ethics: A Quantitative Analysis of the Impact of Unethical Behavior by Publicly Traded Corporations. *Journal of Business Ethics* 16:(5) 537-543 jstor.org/stable/25072921
- Harrison T. 2015. Virtuous reality: moral theory and research into cyber-bullying. *Ethics Inf Technol* 17: 275doi.org/10.1007/s10676-015-9382-9
- Janopaul-Naylor E, Feller E. 2019. Cyberbullying: Harrassment at your fingertips. *Rhode Island Medical Journal*; Providence 102:(9) 7-9 ncbi.nlm.nih.gov/pubmed/?term=31675779
- Kennedy JA, Kray LJ. 2013. Who Is Willing to Sacrifice Ethical Values for Money and Social Status?: Gender Differences in Reactions to Ethical Compromises. *Sage: Social Psychological and Personality Science* 5:(1) 52-59 doi.org/10.1177/1948550613482987
- Koneswaran G, Nierenberg D. 2008. Global Farm Animal Production and Global Warming: Impacting and Mitigating Climate Change. *Environmental Health Perspectives* 116:(5) 578-582doi.org/10.1289/ehp.11034
- Kowalski RM, Limber SP. 2013. Psychological, physical, and academic correlates of cyberbullying and traditional bullying. *Journal of Adolescent Health* 53: S13-S20 doi.org/10.1016/j.jadohealth.2012.09.018
- Latané B, Darley J. 1969. Bystander "Apathy". *American Scientist* 57:(2) 244-268jstor.org/stable/27828530
- Lincoln DJ, Pressley MM, Little T. 1982. Ethical beliefs and personal values of top executives. *Journal of Business Research* 10: 475-487 doi.org/10.1016/0148-2963(82)90006-6
- Macdonald C. 2019. Money, a disconnecting agent: Reminders of money trigger a feeling of disconnection which increases the likelihood of unethical decisions. *The Open Science Journal* 4:(1) 1-10 doi.org/10.23954/osj.v4i1.2200
- Macdonald C. 2020a. Lexicon connection priming and ethics: As our sense of human connection decreases, so too does our ethical decision-making. *The Open Science Journal* 5:(1) 1-7doi.org/10.23954/osj.v5i1.2276
- Macdonald C. 2020b. Graphical connection priming and ethics: As our sense of human connection decreases, so too does our ethical decision-making. *The Open Science Journal* 5:(1) 1-6doi.org/10.23954/osj.v5i1.2275
- Macdonald. 2020c. The Avatar Effect: The harmful consequences of decision-making through a 'separate' entity. *The Open Science Journal* (pending publication)
- Menesini E, Camodeca M. 2008. Shame and guilt as behaviour regulators: Relationships with bullying, victimization and prosocial behaviour. *British Journal of Developmental Psychology* 26:(2) 183-196
- Menesini E, Nocentini A, Camodeca M. 2013. Morality, values, traditional bullying, and cyberbullying in adolescence. *The British Journal of Developmental Psychology* 31:(1) 1-14
- Murphy PR, Smith JE, Daley JM. 1992. Executive Attitudes, Organizational Size and Ethical Issues: Perspectives on a Service Industry. *Journal of Business Ethics* 11:(1) 11-19 jstor.org/stable/25072241
- Peebles E. 2014. Cyberbullying: Hiding behind the screen. *Paediatrics & child health* 19:(10) 527-528doi.org/10.1093/pch/19.10.527
- Perren S, Gutzwiller-Helfenfinger E. 2012. Cyberbullying and traditional bullying in adolescence: Differential roles of moral disengagement, moral emotions, and moral values. *European Journal of Developmental Psychology* 9: 195-209
- Pfungst O. 1965. *Clever Hans* (Rahn CL, Trans.). New York: Holt, Rinehart & Winston
- Pozzoli T, Gini G. 2010. Active defending and passive bystanding behavior in bullying: The role of personal characteristics and perceived peer pressure. *Journal of Abnormal Child Psychology* 38: 815-827
- Rosenberg M. 1979. *Conceiving the self*. New York: Basic Books
- Rosenthal R. 1966. *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts
- Rosenthal R, Rosnow RL. 1969. *Artifact in behavioral research*. New York: Academic Press
- Rosenthal R, Rosnow RL. 1975. *The volunteer subject*. New York: Wiley-Interscience
- Salkind NJ. 2010. *Encyclopedia of research design*. Thousand Oaks, CA: SAGE Publications
- Shalvi S, Gino F, Barkan R, Ayal S. 2015. Self-Serving Justifications: Doing Wrong and Feeling Moral. *Current Directions in Psychological Science* 24: (2) 125-130doi.org/10.1177/0963721414553264
- Sudsakorn C, Rattanawiboonsoom V. 2018. Ethical Business Culture and its Impact on Unethical Behaviors in the Workplace: Conceptual Implications. *PSAKU International Journal of Interdisciplinary Research* 7:(1) 132-140dx.doi.org/10.2139/ssrn.3226468

- Suler J. 2004. The online disinhibition effect. *CyberPsychology and Behavior* 7: 321-326
doi.org/10.1089/1094931041291295
- Soltani B. 2014. The Anatomy of Corporate Fraud: A Comparative Analysis of High Profile American and European Corporate Scandals. *Journal of Business Ethics* 120:(2) 251-274
[jstor.org/stable/42921335](https://www.jstor.org/stable/42921335)
- Valentine ME. 1980. The attenuating influence of gaze upon the bystander intervention effect. *Journal of Social Psychology* 111:(2) 197 psycnet.apa.org/record/1981-25777-001
- Vanderbosch H, Van Cleemput K. 2008. Defining cyberbullying: A qualitative research into the perceptions of youngsters. *CyberPsychology and Behavior* 11: 499-503
doi.org/10.1089/cpb.2007.0042
- Wiid JK, Cant MC, Van Niekerk C. 2013. Moral behaviour and ethical misconduct in Nigerian small businesses. *International Business & Economics Research Journal* 12:(9) 1087-1100
doi.org/10.19030/iber.v12i9.8054
- Ybarra ML, Mitchell KJ. 2004. Online aggressor/targets, aggressors, and targets: A comparison of associated youth characteristics. *Journal of Child Psychology and Psychiatry* 45: 1308-1316
doi.org/10.1111/j.1469-7610.2004.00328.x
- Yuhao L. 2010. The Case Analysis of the Scandal of Enron. *International Journal of Business and Management* 5: 37-41
doi.org/10.5539/ijbm.v5n10p37