Intersect, Vol 16, No 3 (2023)

Critique of Watson and Rayner (1920) and Proposal for Alternative Research: A Look at the Progression of Research and Ethics in Society

Youness Robert-Tahiri
University of Toronto, St. George

There are serious ethical shortcomings within Watson and Rayner (1920). Within it, the authors attempted to condition a fear response into an 11-month-old infant. Specifically, they wanted to verify if they could make an infant scared of an external stimulus, like an animal or an object. Their independent variable was the timing and pairing of a noise with the presentation of a live rat. The dependent variable, fear response, was measured through observations by the authors. Watson and Rayner concluded that they were able to condition a fear response into the child. Before further criticizing their work, its strengths should be highlighted.

Strengths

The strengths of the paper lie in its experimental challenge to non-experimental Freudian concepts, such as erogenous zones influencing personality development. Specifically, Watson and Rayner's experiment seems to highlight the connection between fear, conditioned in this case, and personality; whereas the Freudian concept they refer to seems to negate this and attribute total power to sex for personality development. Although methodological and ethical issues are present in Watson and Rayner, which I will elaborate on shortly, the foundation of their experiment does seem to provide a potential overall effect of fear on personality, which is at odds with Freud's psychosexual development theory. Furthermore, the paper is very easy to read, well-organized, and does a great job of limiting jargon; I believe anyone, regardless of scientific background, could understand the goal of this paper.

Ethical Limitations

Unfortunately, this paper violates multiple ethical principles within the scientific community. To start, there does not appear to be any informed consent from the lone participant or his caregiver. Since Watson and Raynor's work, the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research released the Belmont Report (1979), which addressed issues such as informed consent. Any study that neglects the special attention that vulnerable groups, like

infants, require in experiments, violate the Belmont Report's principle of respect for persons. Moreover, further progress has been made since Watson and Rayner with the release of the American Psychological Association's (APA) Ethical Principles of Psychologists and Code of Conduct (2017); their section on respect for people's rights and dignity being a welcome addition to alleviate issues like Watson and Rayner's lack of informed consent. Furthermore, the Belmont Report's and APA's principle of beneficence were essential additions to the ethical research literature; highlighting the need to avoid causing harm to participants. This prevents experiments engaging in harmful acts like Watson and Rayner's attempt to condition fear in an infant. In addition to the US-based Belmont Report and APA Ethical Principles, there has been progress on the international level within the Declaration of Helsinki (World Medical Association, 2013); with many of their points mapping onto the Belmont Report and the APA Ethical Principles.

Non-Human Animal Considerations

Homing in on animals, section 8.09 of the APA Ethical Principles now requires scientists to ensure the ethical care of any animals used. This was put in place to prevent papers like Watson and Rayner's from omitting information relating to their animal usage. Additionally, the *Guide for the Care and Use of Laboratory Animals* (Institute for Laboratory Animal Research, 2011) was developed so that experimenters disclose their attempts to: replace the animal subjects, refine their experimental procedure to minimize harm, or reduce the amount of animal subjects used. Watson and Raynor were one of many experiments at the time that did not employ anything remotely close to the above described (Franco, 2013). However, the debate between the state of current ethical considerations for animals, or lack thereof, continues to this day (Hajar, 2011).

Theoretical Limitations

In addition to ethical concerns, the author's theoretical framework regarding their participant selection seems questionable. They stated that they are using this particular infant in their study due to his lack of emotionality, and they emphasize how ideal this is for the purpose of conditioning fear. Aside from being ethically questionable, this also goes against later work from Bowlby's (1982) Attachment Theory; where Bowlby suggests that low emotionality could be a sign of avoidance due to poor attachments to caregivers, and *not* of emotional fortitude (as Watson and Rayner alluded to throughout their paper). Moreover, the infant used in the experiment matched Bowlby's criteria of potentially having insecure attachments to caregivers because the infant was in the temporary care of a wet-nurse at the Harriet Lane Home for Invalid Children at the time of the experiment.

Validity and Reliability Limitations

The failure to consider an evidence-based theoretical framework for their rationale of selecting their participant, the infant, may have inhibited the overall internal validity of their experiment. That is to say, maybe what they recorded as a "no fear response" for the infant was actually untrue; maybe the infant was feeling an extreme amount of fear, but due to the infant's avoidant attachment style (the third variable not considered), they were simply not expressing it outwardly. To remedy this, ethical issues aside, a galvanic skin test may have been better suited for operationalization purposes to measure fear that may have been outwardly unobservable. Nonetheless, this lack of internal validity inhibits Watson and Rayner from making any causal claims, which seems to be their aim as described in their final page. They seem to establish covariance (the fear response seems to increase with the conditioning manipulation), and temporal precedence (the conditioning manipulation seems to occur first, and then the fear response), but they are missing internal validity considerations (perhaps another variable, like attachment style, systematically affected their experiment).

In addition to internal validity issues, I have concerns on how the specific observations were conducted for this experiment. There does not appear to be any interrater reliability measures, which hinders the reliability of their results. Moreover, it is unclear how these observations were even coded. As such, observer bias may have affected the conclusions from Watson and Rayner; their expectations could have clouded their interpretation due to their not being a pre-determined coding structure for the observations. This is amplified by their lack of statistical validity, which is absent due to the inability to verify the precision of any estimate. Since the beginning of replication concerns within the scientific community, these previously mentioned issues have been more carefully considered thanks in part to organizations like the Center for Open Science (Center for Open Science, n.d.).

Aside from the ethical issues, this paper's lack of proper theoretical framework regarding participant selection seems to also be problematic. Although they had a decent framework in constructing their hypothesis (i.e., classical conditioning), there seemed to be an overlook in participant selection. Lastly, being an observational study, Watson and Rayner's lack of attention to reliability and statistical measures further hindered their experiment. Overall, their challenge to Freudian concepts through a behaviourist framework was an honourable endeavour, but ultimately, an ethically dangerous one.

Alternative Proposal

Considering the ethical concerns with Watson and Rayner (1920), and the ethical advancements in science since then, a quasi-experiment would be better suited for their research question. Attempting to condition fear in anyone, let alone an infant, violates multiple ethical principles within the

scientific community. Quasi-experiments, on the other hand, do not require an independent variable to be manipulated, therefore there is a reduced chance of ethical violations. For instance, in Watson and Rayner, they attempted to condition fear into an infant, but what if there were already pre-existing groups in society who were already conditioned to fear due to life's challenges? Also, is the specific use of infants even necessary? Therefore, I propose using adult post-traumatic stress disorder (PTSD) patients as the group of study. Individuals who suffer from PTSD already tend to have fear conditioning to external stimuli (American Psychiatric Association, 2013). It seems as if Watson and Rayner were traumatizing the infant by conditioning him to fear and then not reversing it back; however, and unfortunately, there are groups in society, like PTSD patients, who are already suffering in this manner. Therefore, no unethical experimental manipulation would be needed in my proposal.

Participant Considerations

My proposed quasi-experiment involving PTSD patients would answer Watson and Rayner's question by using a low number of participants (i.e., around 4-5). This setup, a small N-design, is helpful in gathering lots of rich information from a small number of individuals. Specifically, I would collaborate with hospitals to ethically coordinate this participant recruitment and, unlike in Watson and Rayner's experiment, informed consent would have to be obtained. Participants would know in advance any risks or benefits, along with being made aware that they could quit the study at any time without consequence. Participants would also be fully debriefed at the conclusion of the study. Moreover, because PTSD patients are a particularly vulnerable group, I would ensure to obtain the approval of an Institutional Review Board to implement any provisions needed before going through with this quasi-experiment. Furthermore, I would ensure that the Belmont Report principle of justice would be adhered to; this requires that the research participants directly benefit from the type of research being conducted (especially if they are a specialized group, which is the case here).

PTSD patients would benefit from my quasi-experiment because, unlike Watson and Rayner, my focus would be in developing an intervention that reconditions fear *out* of the participants. Through self-report, and definitely not through experimental trials, I would ask if they had any particular fears to any external situations, which was the goal of Watson and Rayner's experiment. Since they are PTSD patients, it is very likely that they would be able to list some relevant fears. Afterwards, I would use the same classical conditioning framework as the original experiment, but instead I would reverse-engineer it. That is, instead of pairing the external event with a negative emotion (fear), I would attempt to pair it with a positive one.

Play Therapy

This positive pairing process would revolve around the framework of play therapy, which is an active type of treatment that involves the participant engaging in an activity enjoyable to them (Association for Play Therapy, n.d.). For example, if they loved comedy, I would invite the participant to perform comedic improvisation scenes, which are all made up on the spot, revolving a (at first) very distant theme about their conditioned fear. This would be under the strict supervision of a psychologist on hand, and safety procedures would be implemented in case I notice any harm towards the participant. This would also be in line with the Belmont Report's principle of beneficence, which requires the protection of participants from harm. This means that the play therapy intervention (e.g., comedic improvisation, painting, etc.) would be my quasi-independent variable ("quasi" because there has been no random assignment). My quasi-dependent variable would be the self-reported survey responses regarding the level of conditioned fear involving a specific external stimulus.

Validity Remedies

Although generalizability can be challenging for small-N designs, I still think the data that I would gather would be valuable and applicable to many individuals. My goal would not be to generalize to a large group, but rather to simply help a specific population of PTSD patients. Aside from specifying about external validity, I would ensure that my study's construct validity is adequate by utilizing converging validity means. Specifically, I would ensure that the participants, who are all diagnosed with PTSD, answer other evidence-based questionnaires to measure their level of anxiety and fear, which are prominent parts of PTSD. These converging methods should all align, which would give my quasi-experiment more construct validity by ensuring that I am measuring what I am intending to measure. Furthermore, if the participants engaged in play therapy involving comedic improvisation (for example), comedic improvisation teachers in the community would be consulted to ensure that my proposed comedic interventions have face validity. Lastly, I would ensure that the self-reported questionnaire that measures fear level for participants would use a Likert scale to ensure clarity, and that it would not include any double-barrelled questions to avoid confusing the participants.

Overall, I believe that the ethical and validity revisions in this proposal will allow for further research in fear conditioning. I predict that, and in line with the exposure therapy concept that Watson and Rayner mentioned in their paper, my proposed play-based habituation process would eventually lessen the PTSD patient's conditioned fear response to a specific stimulus. It is my study's aim to empower the participants through play so that they can maybe one day laugh at their pain instead of be frozen by it.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).

Association for Play Therapy. (n.d.). *Play therapy makes a difference*. https://www.a4pt.org/page/PTMakesADifference

Bowlby, J. (1982). Attachment and loss: Retrospect and prospect. *American Journal of Orthopsychiatry, 52*(4), 664–678. https://doi.org/10.1111/j.1939-0025.1982.tb01456.x

Center for Open Science. (n.d.). *Anniversary timeline*. https://www.cos.io/timeline

Franco, N. (2013*).* Animal experiments in biomedical research: A historical perspective*. Animals, 3*(1), 238–273. https://doi.org/10.3390/ani3010238

Hajar R. (2011). Animal testing and medicine. *Heart views: The official journal of the Gulf Heart Association, 12*(1), 42. https://doi.org/10.4103/1995-705X.81548

Institute for Laboratory Animal Research. (2011). *Guide for the care and use of laboratory animals* (8th ed.). The National Academies Press.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. U.S. Department of Health and Human Services. https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3*(1), 1-14. https://doi.org/10.1037/h0069608

World Medical Association. (2013). World Medical Association Declaration of Helsinki ethical principles for medical research involving human subjects. *JAMA: Journal of the American Medical Association, 310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053