

The Monster Study

Turned in on
December 17, 2019

by
Daniel G. Leonard

for the class
ANTH 441 Human Genetics

at
University of Illinois at Urbana-Champaign

taught by
. Karthik Yarlagadda

1 Introduction

In 1939, University of Iowa graduate student Mary Tudor submitted a thesis attempting to tease out the psychological risk factors for the development of stuttering. Studying under Dr. Wendell Johnson at the genesis of speech pathology as a scientific discipline, her research was guided by one of Dr. Johnson's early hypotheses on the origin of the disability. Sixty-two years later, this study was brought to the forefront of the popular press as the "Monster Study," the title of Jim Dyer's (2001) nationally-syndicated column which evoked shock, outrage, and debate, even inspiring a musical composition (Gartman, 2019). In the intervening two decades, scholars have debated the outcomes and implications of the work by Tudor and Johnson.

2 The Study

Johnson suffered from stuttering since the age of five (Reynolds, 2006), which personally affected him greatly and drove much of his career. In the early twentieth century, stuttering was considered innate, an incurable symptom of a broader intellectual disability (Goldfarb, 2006a). More specifically, the leading theory by the 1930s was of incomplete brain hemisphere dominance, in which the right hemisphere is unable to maintain control of the body, leading to stuttering and the (at the time, dreaded) left-handedness (Silverman, 1988). As a successful academic, Johnson rejected this approach, setting his career ambitions to find a cure for stuttering and learn how to prevent it in others (Bloodstein, 2006). Looking inward, he developed his diagnosogenic hypothesis.

2.1 Diagnosogenic Hypothesis

While the theory does not develop until 1959, the first precursors to Johnson's diagnosogenic hypothesis can be traced to earlier works of his. In "Psychological considerations of stuttering" (1936), Johnson explores the sufferer's attitude towards their own condition. He makes two propositions: first, that people who stutter desire to not stutter, and second, that people who stutter expect to stutter just before speaking. Synthesizing these two, he concludes that there exists in stuttering individuals an inhibition of speech that results directly from the fear of negative connotations associated with stuttering. More succinctly, people who stutter do so in part because of a psychological fear of their own expectations to stutter. This self-inhibitory hypothesis also explained stutterers' low socioeconomic position in a way better than that of intellectual disability: "the stutterer is shy and seclusive, and he tends to seek vocations which demand little or no normal speech" (W. Johnson, 1936, p. 23).

To flesh out a more concrete theoretical foundation for the basis of stuttering, Johnson examined the psychiatric diagnosis of childhood stuttering through this self-inhibitory lens. Three main points underlie the methodology and pursuit of Tudor's thesis: first, that the speech of stuttering children at the time of diagnosis is indistinguishable from normal childhood disfluencies (Goldfarb, 2006b); second, the initial diagnosis comes from laypeople, usually parents or teachers (W. Johnson, 1942); third, the child diagnosed as a stutterer begins to experience the aforementioned self-inhibitory phenomenon (W. Johnson, 1936). To confirm or reject this hypothesis, a controlled experiment would be necessary.

2.2 Methodology

Mary Tudor was assigned a project to assess the impact of diagnosis on stuttering children. Johnson sent her to the Soldiers and Sailors Orphans' Home in Davenport, Iowa where there was an experimental population of children. Of 256 orphans, 22 were selected for the study and individually divided into four groups, reproduced in full:

Group IA consisted of five children who had been labelled "stutterers" by members of the institution. An attempt was made to remove the label "stuttering" from the children in this group; that is, they were told that they were not stutterers, but normal speakers who had been erroneously called stutterers.

Group IB consisted of the other five children who had been labelled "stutterers" by members of the institution. In the case of these children the judges endorsed the label.

Group IIA consisted of six normal speakers with varying degrees of fluency. To this group the judges attached the label "stuttering"; that is, they were told that the type of speech interruptions they were having indicated that they were stutterers.

Group IIB consisted of six normal speakers matched in age, sex, intelligence, and fluency with the corresponding six normal speakers in group IIA. No negative evaluative label was attached to this group.

(Tudor, 1939, p. 4)

Over the course of the semester-long study, orphanage instructors would treat the children according to their experimental group rather than to their previous diagnosis. Tudor engaged in regular speech recording sessions with each child, reproducing their speech transcripts in full in her thesis as the data for analysis.

2.3 Results

As reported in the thesis, there was no significant result for individuals in groups IA, IB, or IIB. The sole consequential result Tudor found was in group IIA, the normal speakers told that they stuttered. As Tudor writes in her discussion, "a decrease in verbal output was characteristic of all six subjects; that is, they were reluctant to speak and spoke only when they were urged to" (Tudor, 1939, p. 147).

3 Response

Group IIA was also to be the most controversial aspect of the study upon its public exhumation. While the study had been documented previously in the academic press (see Silverman, 1988), it was an article in the *San Jose Mercury News* that brought controversy and 2006 edited book on the subject.

3.1 From the News Media

Dyer's (2001) news article was syndicated widely and brought great attention to the University of Iowa. In "Ethics and Orphans: The 'Monster Study,'" he describes an unethical experiment which brought permanent hardship to a group of Iowa orphans at the same time as Nazi scientists' similar work across the ocean. The ethical issues with Tudor's research, as described in the article, were many: it was intended from the outset to induce stuttering in children and did so successfully, it used children as subjects, and it did not involve informed consent from all people involved, among others.

Following the article's publication, the six former members of group IIA sued the State of Iowa in 2003, seeking \$13.5 million in compensation for a lifetime of psychological harm. In 2007, the case was settled with the state awarding \$900,000 in compensation to the six participants (News, 2007).

3.2 From Academia

Despite this fanfare, academics have been far more defensive of the study. Ambrose, Nicole Grinager and Yairi (2002) reanalyzed Tudor's data to conclude that Tudor's own conclusion was wrong: "assessments of the two types of quantifiable data, perceptual and speech, clearly indicate that all four experimental questions were answered in the negative" (p. 199). The labeling of group IIA as having decreased in verbal output was unfounded by Tudor's own data, let alone Dyer's deeming them as permanent stutterers. Furthermore, Ambrose, Nicole Grinager and Yairi (2002) find no aspect of the study that would violate ethical norms of 1939, despite their initial reservations that "it is unquestionable that the study was ethically wrong" (Yairi & Ambrose, 2001, p. 17).

Yairi (2006) later went on to criticize Dyer directly, outlining many false and misleading statements in the original *San Jose Mercury News* article. Specifically, the posthumous diagnosis of group IIA children as having become "stutterers" when even Tudor described only mild disfluency, depicting Johnson's diagnosogenic theory as having already been defined rather than in its earliest stages, and painting the article as an exposé when the thesis had been checked out multiple times from its depository in public archives for over half a century.

Regarding the specific ethical questions of Tudor's study, Schmidt, Galletta, and Obler (2006) argue that the most pernicious may be Johnson's imposition of a research project on Tudor, viewing it as a type of professor-student relation that is no longer acceptable in academic culture. Asking whether such a study would pass an IRB today, Schwartz (2006) contends that the Tudor study, as conducted, would not, for the reason that negative-feedback experiments are rarely acceptable in today's research environment.

In contrast, Harris (2006) refutes the criticism of Johnson having performed the study as the test of an unexplored hypothesis, listing several seminal works in the field which were only possible because they were exploratory of new ideas. Most radically, N. Johnson (2006) defends the study outright, explaining several postulates: harm was neither intended nor done, as there is no evidence for permanent repercussions; children were the only acceptable subject population for the hypothesis, being based on developmental speech pathology; informed consent in this case was provided, as although deception of the children was required the administrator of the orphanage provided consent; the experiment was limited in scope and time, being limited to only a few children and lasting only one semester; finally, there was adequate post-study care for the subjects. In a departure from his peers, he argues the Tudor study would not only pass IRB approval in today's academic environment, but that it is radically safer than much contemporary research.

4 Discussion

Personally, I am convinced by Nicholas Johnson's defense of Tudor's study. Although its methodological flaws outlined by Ambrose, Nicole Grinager and Yairi (2002) preclude the study's use as empirical science, I find no ethical flaws in the methodology or practice of the study, especially in the context of 1939. In my own reading of Tudor's thesis, I see a graduate student who performed an experiment with all the care and nuance human-subjects research requires. As an anthropology student, the study is notable to me in that it took great care to describe the thoughts and feelings of children with and without the "stutterer" label. A great departure from the usual early-20th-century scholarship of disability as an innate flaw, Tudor's work reads almost as an ethnography of identity and disability. The only flaw I find in the thesis is the glaring lack of a fully fleshed-out theoretical foundation, as the introduction section is startlingly brief. Whether this was an artifact of its time I am not sure, but a more grounded introduction would surely be necessary in contemporary research.

Save for the settlement from the State of Iowa, the study had no impact on ethical guidelines of today. Being already more than a half-century old at the time of the media frenzy, IRB guidelines already existed in very well-defined terms, and there was no change to how they operated as a direct result.

References

- Ambrose, Nicole Grinager, & Yairi, E. (2002). The Tudor study: Data and ethics. *American Journal of Speech-Language Pathology*, 11(2), 190–203. doi:10.1044/1058-0360(2002/018)
- Bloodstein, O. (2006). Research in stuttering at the University of Iowa circa 1939. In R. Goldfarb (Ed.), *Ethics: A case study from fluency* (Chap. 3, pp. 27–34). San Diego: Plural Publishing.
- Dyer, J. (2001). Ethics and orphans: The ‘monster study’. (2001), 1A.
- Gartman, E. R. (2019). Monster study. Re-mix it competition, university of illinois at urbana-champaign. Retrieved from <https://hdl.handle.net/2142/103717>
- Goldfarb, R. (2006a). An atheoretical discipline. In R. Goldfarb (Ed.), *Ethics: A case study from fluency* (Chap. 8, pp. 117–138). San Diego: Plural Publishing.
- Goldfarb, R. (2006b). Diagnosis. In R. Goldfarb (Ed.), *Ethics: A case study from fluency* (Chap. 2, pp. 13–26). San Diego: Plural Publishing.
- Harris, K. S. (2006). Some physiological studies on stuttering. In R. Goldfarb (Ed.), *Ethics: A case study from fluency* (Chap. 7, pp. 97–116). San Diego: Plural Publishing.
- Johnson, N. (2006). Retroactive ethical judgements and human studies research: The 1939 Tudor Study in context. In R. Goldfarb (Ed.), *Ethics: A case study from fluency* (Chap. 9, pp. 139–200). San Diego: Plural Publishing.
- Johnson, W. (1936). Psychological considerations of stuttering. *Exceptional Children*, 3(1), 22–24. doi:10.1177/001440293600300107
- Johnson, W. (1942). A study of the onset and development of stuttering. *Journal of Speech Disorders*, 7(3), 251–257. doi:10.1044/jshd.0703.251
- News, B. (2007). Huge payout in US stuttering case. Retrieved from <http://news.bbc.co.uk/2/hi/americas/6952446.stm>
- Reynolds, G. (2006). The stuttering doctor’s “monster study”. In R. Goldfarb (Ed.), *Ethics: A case study from fluency* (Chap. 1, pp. 1–12). San Diego: Plural Publishing.
- Schmidt, B., Galletta, E., & Obler, L. K. (2006). Teaching research ethics in communication disorders problems. In R. Goldfarb (Ed.), *Ethics: A case study from fluency* (Chap. 5, pp. 63–82). San Diego: Plural Publishing.
- Schwartz, R. G. (2006). Would today’s IRB approve the Tudor Study? ethical considerations in conducting research involving children with communication disorders. In R. Goldfarb (Ed.), *Ethics: A case study from fluency* (Chap. 6, pp. 83–96). San Diego: Plural Publishing.
- Silverman, F. H. (1988). The “monster” study. *Journal of Fluency Disorders*, 13(3), 225–231. doi:10.1016/0094-730X(88)90049-6
- Tudor, M. (1939). *An experimental study of the effect of evaluative labeling of speech fluency* (Master’s thesis, State University of Iowa). doi:10.17077/etd.9z9lxfgn
- Yairi, E. (2006). The Tudor Study and Wendell Johnson. In R. Goldfarb (Ed.), *Ethics: A case study from fluency* (Chap. 4, pp. 35–62). San Diego: Plural Publishing.
- Yairi, E., & Ambrose, N. G. (2001). The Tudor experiment and Wendell Johnson: Science and ethics reexamined. *ASHA Leader*, 6(13), 17.