

Chapter 5

Carl Stumpf and Control Groups



Julia Kursell

In the fall of 1917, a group of students visited the Institute of Psychology at the University of Berlin. During their lectures in psychology, they had been invited to participate in an experiment. Explanations were provided on the spot. The students had to enter a booth mounted in one of the Institute's rooms, where they found the opening of a tube connecting the booth to the adjacent room. Through that tube they would hear sounds, which they were supposed to judge. Did they recognize any vowel? And what did they think of the vowel's quality? Such were the questions they received in advance.

In a paper read to the Prussian Academy of Sciences in 1918, the head of the Institute, philosopher Carl Stumpf, commented on the prior knowledge—or rather ignorance—he sought in these experimental subjects. He also pointed to the difficulty of judging sound without previous information:

They had no idea about the entire setup and its purpose. They were only told that they would hear vowels. Because a sound so short and without the characteristic beginning [that was cut off from the transmission] is so ambiguous, such previous information is necessary to make any interpretation possible. (1918, 353)¹

The purpose of the experiment was to test the sound quality of synthetically produced vowel sounds. With a gigantic structure—the so-called interference device (*Interferenzeinrichtung*) for sound analysis and synthesis—occupying almost all of the rooms of the Institute, it had become possible to emulate the sound spectra of

The author wishes to thank the editors of this volume, all participants of the workshop and especially the two readers Christoph Hoffmann and Klodian Coko for their support and insightful commentaries.

¹ Translations are mine unless otherwise stated.

J. Kursell (✉)
Faculty of Humanities, University of Amsterdam, Amsterdam, Netherlands
e-mail: J.J.E.Kursell@uva.nl

vowels so convincingly that uninformed subjects were likely to recognize them. This was exactly the role of the visitors: they were test subjects in what Stumpf called “uninformed experiments” (*unwissentliche Versuche*). An Institute staff member prepared the synthetic production of the vowel under scrutiny in accordance with previously determined data, and then sent it to the booth in random alternation with the sound of a singer in another room. The students’ answers helped the researchers determine whether the synthetic vowels withstood a comparison. Stumpf explained the rationale of this quality check:

Due to the tendency to think of the synthetic vowels as reaching truth to nature as soon as a slight resemblance has been achieved, I worked not only with alternating observers, but I also systematically carried out uninformed experiments the statistics of which I compiled. (Stumpf 1918, 353)

As it turned out, the uninformed subjects did not reject the quality of the artificial vowels as deficient. On the contrary, they often found them more convincing than the ones produced naturally. More interesting for the present chapter, though, is the role of the uninformed experiment in Stumpf’s methodology. As the quotation above reveals, Stumpf used a comparison between the presence and absence of a concrete condition of judgment, which allowed him to control the observers’ bias in new ways. Whether or not the goal of synthesizing vowel sounds had been reached could be determined only through subjects who were ignorant of that goal.

The experiment with uninformed subjects was part of a setup that involved control on several levels. First, the available data about frequency components in vowels were compared with data extracted with a new device for sound analysis. These data were then recreated synthetically and compared with the original vowels by trained observers. Finally, the uninformed subjects were exposed to the comparison between synthetically and naturally produced vowels. In all steps, the comparison between independently produced sets of data was central. This is in line with the etymology of “control” tracing to the French “contre rôle” or counter roll: a second, independent list to be compared with a first. The term emerged in processes of administration and was soon used in the context of scientific experimentation. Similar to what historian of psychology Edwin G. Boring (1954) states for the English word “control,” its German equivalent gained currency in the first half of the nineteenth century.²

German writer Johann Wolfgang Goethe, for instance, used the word for administrative matters and made the character Odoard in his novel *Wilhelm Meisters Wanderjahre* (1829) do so too when explaining the supervision needed to instigate an agrarian reform.³ The word is not mentioned in the dictionary of the German

²Boring also mentions the use in German psychophysics. For a close reading of this paper see Schickore (2019).

³This is in line with the emphasis on administration in the entry “Kontrolle” of the term’s history given in the German dictionary *Meyers Großes Konversationslexikon* from 1905 to 1909. This entry draws back to French uses of the seventeenth century and mentions that the German equivalents “Gegenschreiber” (counter-writer) and “Gegenbuch” (counterbook), as well as the direct borrowing “Kontrolleur,” appeared in mining contexts “a long time ago.” See *Meyers Großes*

language initiated in 1838 by the two brothers Johann Jacob and Wilhelm Grimm (*Deutsches Wörterbuch* 2021), even though that project grew to comprise sixteen volumes during 123 years of collecting and editing. However, by the time Stumpf published his first book on the psychological origin of spatial representation (Stumpf 1873), both the verb *controllieren* (to control) and the noun *Kontrolle* were established. In later publications, Stumpf used the terms more and more frequently, consistently referring to instances of comparison by which the calibration of experimental setups could be checked and the validity of findings confirmed.

In Stumpf's vowel experiments, several functions of control, as well as several strands of its history, intersect. The present chapter will discuss them under the umbrella of the term "control group." Control groups delineate the process of experimentation in experiments with human subjects. They constitute what the "other things" are, the *ceteris paribus* that are supposed to remain unchanged when the main group in the experiment undergoes a certain intervention. Control groups or "unbiased comparison groups" (Chalmers 2001) can consist of subjects who do not know about the aims of the experiment; they help to conceal from those performing the intervention on whom they perform it, thereby counteracting bias; and they often consist of randomly chosen subjects, thereby counteracting bias in the researchers who might otherwise privilege a certain group without noticing. All of these functions help to create a group of subjects that remains "unchanged" in comparison with the experimental group through measures of blinding and randomizing.

In the notion of the control group as instantiating a gauging standard opposed to the experimental group, historian Trudy Dehue (2005) has identified two important assumptions. These assumptions have been taken for granted in the contemporary notion of the control group, but came about only gradually during the nineteenth century. First, the groups had to be understood not as consisting of individuals but as representing "populations," so as to—the second assumption—make them susceptible to statistical treatment. In addition, this chapter points to yet another genealogy of control groups, namely in an experimental logic pairing two states, the positive and negative, of the same condition. To reduce variation this way Stumpf needed neither a notion of population nor the law of large numbers. This chapter will discuss how he devised his method, conceiving of himself as a philosopher who understood the psychology he contributed to as a subdiscipline of philosophy.⁴

Just as much as the vowel study required control, the comparison of the subjects' judgments produced further insight into Stumpf's other, perhaps main, subject: the study of judgment itself. Starting out as a philosopher who also integrated collections of individual judgments, his method became experimental and eventually led him to seek judgments more and more systematically. During this search he refined his theory of judgment and explicated it more fully.

Konversationslexikon (6th edition of 1905–1909), digitalized in "Wörterbuchnetz des Trier Center for Digital Humanities," version 01/23, <<https://www.woerterbuchnetz.de/Meyers>>, accessed 03 Sept. 2023.

⁴On this, see, for instance, Kaiser-el-Safti (2011) and Martinelli (2015).

The work on phonetics can be seen as a culmination of this development. Stumpf's ingenuity in devising measures of control (*Kontrollmaßnahmen*) involved practical matters in unique ways, allowing me to ask how the fleeting nature of sound prompted functions of control within his research on aural judgment. To unfold the peculiar way in which Stumpf's experimental practice combines control and judgment with his own understanding of what logic should be, this chapter proceeds in three steps, corresponding to three sections of this paper. The first section introduces the experimental setup of the vowel study. It has two parts: the first introduces the workings of the setup, the second focuses on the experiment with uninformed subjects. The second section reviews how Stumpf's method of comparing judgments evolved from his first experiments on auditory judgment after 1873. This section partly confirms the findings of Dehue, while also showing the way toward his theory of judgment. The third section discusses that theory. These three steps will help me to discuss the notion of a control group as an operation rather than as a term. By this means I hope to contribute to the project of researching the history of what becomes a term of art at a given moment in time.

5.1 The Interference Device

5.1.1 *Measures of Control for Acoustic Experimentation*

In 1926 Stumpf published a book titled *Speech Sounds: Experimental-Phonetic Studies with an Appendix on Instrumental Sounds (Die Sprachlaute: experimentell-phonetische Untersuchungen nebst einem Anhang zu Instrumentalklängen)*. The book's main part detailed the work with the interference device. When its construction began in 1913, the device's scale was unprecedented in acoustic research. It comprised two independent systems of tubes, one serving to analyze and the other to synthesize sound. The operative principle was interference: from actual sound waves propagating through the tubes, single frequencies were subtracted by adding vertical spikes to the main tube. The length of the spikes was calculated so as to project the reverse pattern of rarefaction and compression onto the partial wave in question, thus canceling out that frequency component in the overall sound. Potentially, all frequency components could be canceled from the incoming sound with all spikes added to the main tube. The spikes could also be inserted separately, enabling the researchers to generate various configurations to test.

The other part of the structure, the synthesis system, also used interference. There, periodic sound was purified by interference as described above, so as to obtain simple tones consisting of a single frequency. From these simple tones the "synthetic" sound was composed. For instance, the pattern of frequencies resulting from the analysis could be recreated. For these, the simple tones resulting from the purification were joined into a single tube at a place best understood as the device's control room. Both parts of the device provided a spot for an observer in this room. The synthesis structure also allowed some limited manipulation of the incoming

simple tones; tones could be selected and their intensity changed by means of mechanical devices. The tones could be dampened or fully eclipsed with the help of clamps around rubber fittings that were fixed to the ends of each tube. The rubber fittings eventually merged into a single tube, so as to propagate the recreated component pattern to a spot of observation or the booth where the uninformed experimental subjects were located.

“The more finely a method of investigation operates, the more complicated the devices used must be,” Stumpf wrote, when explaining the needs of acoustic experimentation in the introduction to his book (1926, 8). His own technical setup certainly met this criterion. Although the interference device did not yet involve electrical transduction of the sounds, operating instead on acoustic sound propagating through the system, it opened up many new procedures that would become standard features when psychoacoustics labs of the interwar period began using electronic technology on a large scale. One feature of particular interest here is connected to the dimensions of the Berlin device. The tubes propagated actual sound waves, the size of which, for human hearing, ranges between two meters and some millimeters. A plotted floor plan Stumpf added to his publications shows the setup (Fig. 5.1). The entry of natural sound into the system happened at a distance, in different rooms. No sound source was located in the control room, and there the observer could only listen to the tubes’ outputs. For the first time, seeing and hearing were systematically disconnected.

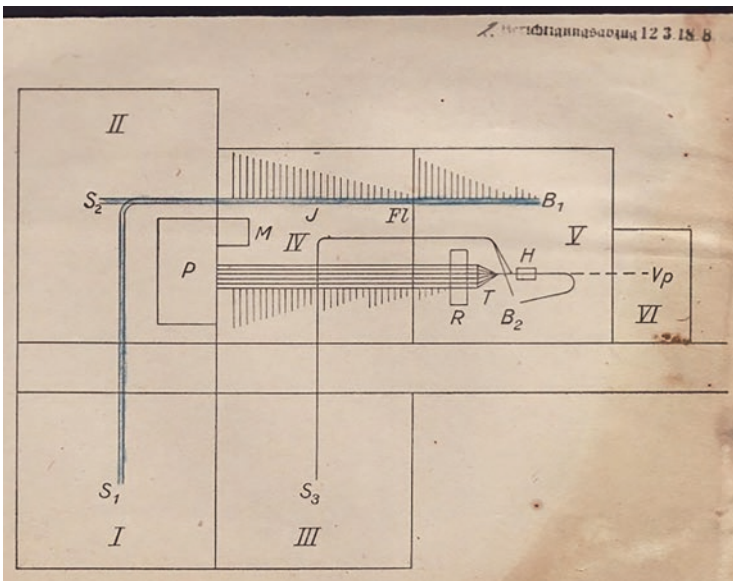


Fig. 5.1 Plotted schema of the interference apparatus of the Institute of Psychology, here taken from a blueprint for Stumpf (1918). The schema was used again in Stumpf (1926) on p. 44. Stumpf Papers, with the kind permission of the Ethnologisches Museum, Staatliche Museen zu Berlin

While earlier experiments attempting to do the same, such as Stumpf's own experiments from 1910 on recognizing the sound of musical instruments from an adjacent room (Kursell 2013), remained incidental, that disconnection became a basic feature in the new device. This feature included the booth for placing uninformed subjects. The method to investigate sound thereby became not only more fine-grained and more complicated, but also more stable, while leaving sound to auditory observation alone. As we shall see, Stumpf took this disconnection to enable also the disconnection of previous knowledge and observation.

The method of interference itself raised the quality of the data on vowel sounds to such a level that the sounds became an object of interest in their own right. In this regard, the work built on acoustic experimentation as it had been introduced into the field by Hermann von Helmholtz. Helmholtz was the first to assume that the ear analyzes sound by reacting to the frequency components selectively. Although he could not determine which mechanism exactly was responsible for reacting only to the frequencies present in a sound—his own “resonance theory” was proven wrong by György Békésy 50 years later—he did everything he could to test the usefulness of this hypothesis.⁵ He was also the first to build an apparatus “for the artificial construction of vowels,” often referred to as the first synthesizer. This instrument allowed him to re-instantiate frequency patterns he had determined before, using sets of resonators—hollow spheres with two openings that would react to a single frequency and were held to the ear, while listening to, e.g., a sung vowel. His synthesizer provided only eight and later twelve frequencies to choose from, and the resemblance to actual vowels was weak.⁶ Yet, with the help of a keyboard allowing him to manipulate the strength of each frequency separately and in quick succession, he could enhance the slight differences in the sound of the patterns. A trained pianist, he could change between frequency patterns quickly and distinctly. A minimal distinction could thus be claimed, which was sufficient to confirm that the ear somehow in fact discriminated the frequency patterns in question.

Although a resemblance to actual vowels was not strictly necessary for Helmholtz's claim, it greatly helped the rhetoric: no other sound could be described in written text so easily for so large a community. Stumpf himself fully assented to it. He summarized the history of vowel synthesis in his book, mentioning many testimonies of experimenters who did not manage to obtain convincing vowel sounds with replicas of Helmholtz's apparatus. He corresponded with physicist Felix Auerbach, who reported that he recognized a vowel only occasionally when configuring the apparatus in order to finely set the required values (1926, 167). While Stumpf held a position in Munich before coming to Berlin, he had access to Helmholtz's original tuning forks, then stored at Deutsches Museum. He could not

⁵After scarce attention since the call in Hiebert and Hiebert (1994) to address music in Helmholtz's work, the literature on Helmholtz's acoustic research has been vast during the past 20 years. For general accounts, see Steege (2012) and Kursell (2018) and for more specific questions, see Jackson (2006), Pantalony (2009), Hui (2013), and Hiebert (2014). Among earlier studies, Scherer (1989) is notable.

⁶On this, see Pantalony (2005, 2009).

use them, however, as one central element, the interrupter-fork, was missing. This did not shake his assumption that Helmholtz did hear vowels: “After all: a Helmholtz cannot be lured with fancies” (168).

Stumpf’s own vowel studies no longer had the function of the *experimentum crucis* that was to decide whether the ear can be said to “analyze” sound. If, for the testing of a hypothesis, it was sufficient that the sound leaned toward a noticeable distinction, Stumpf read the description to the letter and embedded its two components—analysis and synthesis—into a rigid experimental architecture with new points of fixity and openness. Analysis was delegated to the interference apparatus, which provided data about vowels. The synthesis, as in Helmholtz, was supposed to re-instantiate them, but in this new method the re-instantiation would serve to control the data quality and not to test the validity of the connection of analysis and synthesis as such. The analyzing ear was taken for granted in Stumpf’s setup. The focus then moved to the mind.

Indeed, the mind was at stake in Stumpf’s systematic manipulation of prior knowledge. The interference device offered multiple possibilities to situate human observers, but these observers also played a crucial role in making the device function. They monitored change in the analysis structure and determined the strength of the purified tones to be combined in the synthesis structure. They also verified quality in the uninformed experiments. The obvious reason for this human factor in controlling experimentation was that Stumpf could not measure intensity. Only with a concept of sound as energy could amplitude be measured, but such a concept came about only with electroacoustics (Wittje 2016). In Stumpf’s apparatus, all the researchers could do was estimate the strength of a component. The ear remained the judge in matters of acoustics, as Stumpf never stopped insisting.

The two spots for observers in the control room shown on the floor plan (Fig. 5.1, room V) indicate two modes of judging sound: observation of a process, and comparison among results. In the analysis structure, sound could enter at three points: room I or II, with S indicating the position of a singer, and at a third point for whispered vowels (*Flüstervokale*), indicated in room IV with the letters Fl. As the use of interference depended on sound that could be kept constant for a certain amount of time, sung vowels worked the best for this setup. In rooms IV and V, then, the actual process of canceling out frequency components took place. For this procedure the spikes were opened one by one. The observer, located at point B1, monitored the change in sound and its overall intensity. After the sound had disappeared, the procedure was reversed and the sound rebuilt, now closing the spikes one by one, until the sound was transmitted unchanged through the tube again. Once again, the observer’s task was to monitor the change.

On the other tube ending in the control room, marked B2, the observer had to take another action. This spot was connected to the synthesis structure initiating at a soundproof box (P) in which organ pipes were mounted. The pipes were driven by a motor in yet another soundproof box (M). Along the way through room IV, the pipes’ sounds were the purified from overtones (all frequency components except the lowest or fundamental frequency). They then entered room V as simple tones. Here the sound intensity was regulated (R). The observer handled the clamps around

the tubes' rubber fittings while monitoring the resulting sound changes. All components merged at T. The observer then had several options to induce comparison by choosing to listen to several tube endings. One ending transmitted the sound from room III, where a singer could be signaled to start producing the sound in question with an electric bell. Both the synthetic sounds and the singer's sound were brought to a switch that enabled the observer to choose between them to either listen to them him- or herself or to send them along to the booth.⁷ There, the uninformed subject could hear them without suspecting their twofold origin.

5.1.2 *Functions of Control in the Interference Experiments*

From a table-top experiment with Helmholtz, the comparison between analysis and synthesis turned into the content of an entire Institute with Stumpf. Vowels were no longer supporting the rhetoric, becoming instead the object of analysis. Analysis and synthesis, in turn, could be carried out and observed in much greater detail, using the new procedure of the step-by-step canceling or adding of single frequency components with the help of the interference structure. Control was at the core of the structure's division into the two independent systems of tubes for sound analysis and synthesis. The division provided the researchers with corresponding sets of data for comparison. That comparison, however, could not skip the human ear, the final judge for whether a sound could be considered a vowel or not. Control was thus not encapsulated in an exchange between B1 and B2, but spilled over into other points in the setup as well.

Two basic categories of control stand out: the technology and the human observer. They prompted different regimes of control. On the one hand, the fine-grained analysis that was so important to Stumpf required a constant monitoring of the setup's functioning. Thus, *Die Sprachlaute* discusses technical problems at great length. The tubes distorted the sound, to begin with; this could be partly remedied using funnels for the singer, but the funnels had their own impact on the sound. The sound itself could not be controlled with the ear alone. Additional tools were needed, because even when below the threshold of hearing, a sound might nevertheless distort the devices' functioning. This was true, for instance, for the presence of unwanted components in the allegedly pure sounds used in the synthesis. Tuning forks with frequencies that deviated slightly from those of unwanted components were held in front of the openings so as to make them audible as beats in the forks' audible sound. Finally, the basic principle of the structure, sound canceling by interference, was difficult to handle. It could have side effects, such as the canceling of higher frequencies that fit into the same wave pattern or slightly deviating frequencies being reduced below the threshold of hearing within a certain range.

⁷Stumpf had female assistants and doctoral students at that time, such as Katharina von Maltzew, who later went to Soviet Russia.

On the other hand, the observer could not be trusted. Stumpf reports, again in great detail, about their failures, including his own. He writes about surprising observers by interpolating sounds that had not been agreed on beforehand into a series of tested items, such as a consonant in a series of vowels. He constantly compares humans and devices using the same term “Einstellung” for both. One observer, for instance, stubbornly “recognized A” when a “whispered Ö” was spoken into the entry point for whispered vowels, although Stumpf assures the reader that the “Ö” by then was fully recreated in a process of building the components with the analysis device. He also expresses his amazement about “untrained” observers, such as a group of students and staff from the university department of modern languages, to whom he introduced the workings of the setup:

One day, I demonstrated the change of vowel sounds with interference tubes to a group of members from a seminar for modern languages, among them a lecturer. The vowel Ö was being deconstructed, and long after it had transformed into a pure and even dark-shaded O, the first observer [from this group], a lady, insisted in still hearing Ö. This assessment was taken up by all ensuing observers, who had heard her assessment. I almost began to doubt my own ears, until a reliable staff member, Dr. Wertheimer, was called and immediately and without previous information recognized O. (Stumpf 1926, 51)

The cameo of Gestalt psychologist-to-be Max Wertheimer is an aside in this anecdote about the distinction of trained versus untrained observers. At the time, Wertheimer was working with Stumpf, and Stumpf praised his fine ear. However, even the best observers kept failing in a specific task:

There is one point at which even the most trained observer is exposed to a constant psychological influence: the results of decomposition and re-composition consistently deviate, as the stages of transformation are situated at a lower point [of the acoustic spectrum] during re-composition than during decomposition. (Stumpf 1926, 51)

Starting from the fully present transmission, the observers were ready to note any small change, whereas the opposite direction—the re-composing of the vowel from its lowest partials—prompted them to recognize a reappearance of the vowel at the earliest moment. As a result, recognition was lost and gained at different points in the two directions of the process. The reaction of human hearing to language was not like a measuring device, but rather was sensitive to immediate context.

Stumpf’s comment on this deviation demonstrates that his psychological interests were not absent while generating data for phonetics. “That difference can only have psychological causes,” he noted, drawing an analogy to the difference in the threshold of audibility when a sound source was moved toward or away from the ear. He explained the deviating points of loss and recovery of the vowel’s “specific character” in the same way. “To this diverging behavior,” he concluded, “one is submitted even with a high degree of training and even when de- and reconstruction succeed each other immediately” (52). As these observations demonstrate, the measures of control generated their own surfeit of research findings.

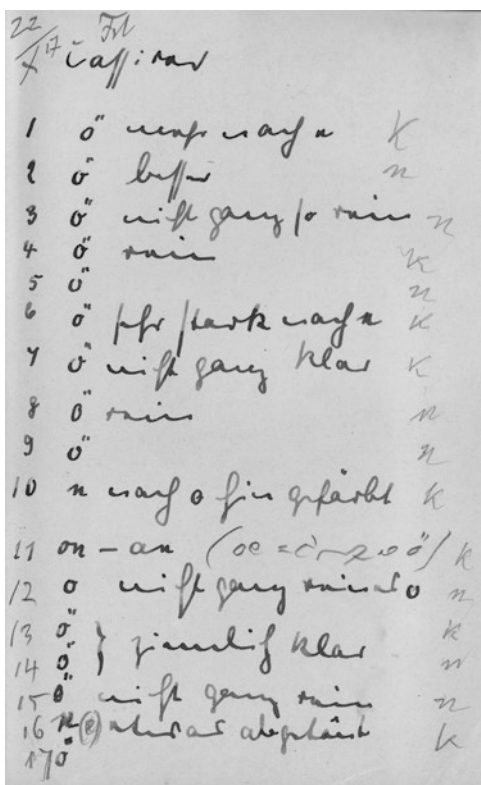
The experiment with the uninformed subjects can be seen as a counterpart to the observer comparing natural and synthetic sounds. When recounting it in the book, Stumpf added a detailed description of the procedure, beginning from what has

remained the practice in psychological experimentation ever since: “To an invitation during my psychology lecture to participate as observer in my vowel studies, 30 students, both female and male, reacted” (182). He first tested these students with regard to their general ability to recognize any sound in the transmission, using only vowels produced by a singer. Of these, eighteen succeeded and were invited to the actual tests. Each vowel was tested with five series of ten pairs of vowels, one natural and one synthetic, in predetermined but randomly chosen sequences. The experimental subjects were instructed as follows:

You will hear vowels of very short duration. Ask yourself when you hear the first of them, which vowel it is and whether its transmission is good or deficient, and if the latter, in which regard, e.g., E too much towards Ä. When you hear the second, ask yourself whether it is the same and if so, whether it sounds better or less good than the one before and why. Then you will always hear pairs that you should compare. Anything remarkable should be noted. (Stumpf 1926, 183)

Although the subjects most often did not follow the instruction to compare pairs, Stumpf found the results to be sufficient for his purpose. Figure 5.2 shows the notes by one of the subjects from October 22, 1917. Reacting to samples of the vowel “Ö,” “Fräulein Cassirer” wrote down on the left side which vowel she thought she had heard. The experimenter added on the right with the letters “k” (*künstlich*, “artificial”)

Fig. 5.2 Notes by a test subject (ink) in an “uninformed experiments” on the vowel “Ö”, comments in pencil by Carl Stumpf. Stumpf Papers, with the kind permission of the Ethnologisches Museum, Staatliche Museen zu Berlin



and “n” (*natürlich*, “natural”) whether the sample was naturally or artificially produced, also providing other necessary information for laboratory purposes, such as the date, numbers where she forgot them, or which vowel she meant when her handwriting was bad. In between, one reads Miss Cassirer’s comments, such as “more towards e,” “better,” “not fully clear,” and “pure,” all conveying her estimate about the quality and distortions in what she discerned. These remarks were fully in line with what Stumpf had asked for: “anything remarkable should be noted.”

Stumpf was pleased with the outcomes of the experiment with the eighteen students. In his usual way, he commented with a subjective tint:

Often the exp. subjects stated somewhat depressed at the end of a series of samples that they had not found any significant differences, that they had always heard the same vowel, which I took note of not without some hidden pleasure. From all their comments it was clear that also during the experiment the subjects had no clue that natural and artificial vowels were presented alternately. (Stumpf 1926, 183)

The experiment confirmed the expectation that the synthesis plausibly reproduced vowel sounds, based on the data generated in the analysis. It also introduced new methodological components, such as blinded testing, random samples, and a statistically relevant number of answers to avoid individual bias:

Individual propensities were showing up here as well. For one subject no A, whether natural or synthetic, was bright enough. Especially regarding A, the expectations indeed differ considerably among individuals. Another subject always heard the natural E to be closer to \ddot{A} , which might in fact be objectively not without a reason. It is exactly because of such small individual differences that a larger number of subjects had been involved. (Stumpf 1926, 183)

To sum up the explanation of how control guided phonetic experimentation here, we can say that the ephemerality of sound, the inability to measure amplitude, and the subjectivity of auditory observation were tackled with a triangle of control instances: first, by monitoring the manipulations in the analysis device in deconstruction and reconstruction; second, by comparing the resulting data with their re-instantiations with the synthesis device; and third, by presenting the synthetically recreated sounds to the uninformed subjects.

The interference device embodied strategies of granting independence for the gathering and comparing of data in its architecture. It provides rich insights for disentangling functions of control, as they were discussed by Jutta Schickore for the life sciences around 1800 (2021a, b). While such a close analysis could instantiate what is subsumed under notions such as Hans-Jörg Rheinberger’s “technical object” (1997, 2023), it is important to note that audition, to some extent, placed the emphasis not on the counterpart notion of the epistemic thing, which describes the moment when the research object, from hindsight, can be understood to have guided the process of experimentation. Instead, Stumpf’s experiments on audition dealt with the defaults that subjects fell back on when placed in a situation of ignorance. For disentangling the distributed action in terms of functions of control, it is therefore important to expand the analysis to the ways in which this research took such defaults into consideration. This takes the chapter back to Stumpf’s earlier work on the psychology of auditory perception and cognition.

5.2 Comparing Judgments

In the triangle of control instances, the experiment with uninformed subjects presents a region of overlap between two competing interests. The subjects acted, as Stumpf writes, like rabbits and frogs, whom the experimenter does not query how they experience what is done to them. However, the opposite occurred in this uninformed experiment: the subjects were asked to take down their experience. The commentaries are literally inserted between the columns of data, pointing to another area of Stumpf's interest: the study of judgment.

He had pursued this study since his first appointment as professor of philosophy at Würzburg University, where he succeeded his former teacher Franz Brentano in 1873. At Würzburg he began researching the psychology of auditory perception and cognition or, in his own terms, tone psychology. He would later publish two volumes with the title *Tonpsychologie*. The first volume, appearing in 1883, dealt with sensory judgment more generally and the judgment of single and successively heard tones more specifically. The second, from 1890, discussed the judgments of two tones given simultaneously and the theory of fusion that Stumpf would remain known for. The two volumes together grew Stumpf's reputation as an experimental psychologist. His renown procured him further positions as professor of philosophy, first in Prague, then at Halle and Munich Universities, and eventually in Berlin, where he took up the position of chair of philosophy in 1894.

Back in the mid-1870s, Stumpf's interest in the general reliability of tone judgments prompted him to invite people to his home who said they had no talent for music. He described this endeavor in the first volume on tone psychology as follows:

At first only with the intention of getting a more definite idea of the degrees of unreliability that occur in judgements about tones, years ago when I was in Würzburg, [...], I asked several people – otherwise well-educated and normal in hearing, but very unmusical – about their judgement as to which of two tones is higher. These people were: Miss C., completely unmusical according to her own statement and those around her; Dr. K., who assured me that he has no clue about music; W., man of private means, who is not disposed to music and ignorant about it; S., man of private means who is, according to his own statement, able to retain easy melodies, but hostile to the violin, and almost never engaged in music in his youth; finally, the students Be. and Bo. I preferred the question “which tone is higher?” to that of “equal or different?”, for I believed this would give me insight into the general conditions of the qualitative judgement. (Stumpf 2020 [1883], 201, translation slightly modified after Stumpf (1883)).

The reliability—or, as phrased here, the unreliability—of judgments was the object of this experiment, and the experiment later developed into a full-fledged method. What Stumpf calls “conditions of judgment” (*Urteilsbedingungen*) could be manipulated by contrasting two complementary conditions: the judgments of those who do, as opposed to those who do not, have a specific and well-defined predisposition, precondition, or, as he would say in *Die Sprachlaute*, “setting” (*Einstellung*) for

making a judgment.⁸ The question which of two notes was higher, as opposed to that about just noticeable differences, targeted the subjects' ability to find their way in the system of Western tonal music. To grasp pitch in that type of music means to subsume the spectrum of a periodic sound under one value and to understand this value as a tone or note that can be situated within a scale. For this one must be able to align the values on the basis of the parameter of "height," which the English language subsumes under the concept of pitch, but which in German appears in the compound *Tonhöhe*. A listener who could not grasp pitch in this specific sense, even if sensitive to sounds or tones being different among themselves, would not understand the rules of tonal harmony and counterpoint. Later musicality tests would continue to use this question.⁹

It is important to note that the judgments of the unmusical were interesting exactly because they were true without being correct. These subjects opened the possibility of working with false judgment in controlled ways, namely as inside or outside a conventional symbolic system. They allowed Stumpf to draw the distinction not in the physicality of the subjects' hearing, but in their access to a particular and very specific set of rules whose application also relies on a subject's exposure and training. While the answers of the unmusical subjects were perhaps not random, they did not match with a particular system in which they, for whatever reason, did not participate. But that system was also not the experiment's main interest, because the intent was to discover the extent to which subjects participate in and have access to any such system. In fact, Stumpf later changed the object of investigation, but he always searched for what he called "psychic functions" at play in accessing these systems, and confronted subjects with tasks that presupposed access. The first task Stumpf explored systematically was the judging of simultaneous sounds. Later he attempted to find out whether the confrontation with musical systems other than Western tonal music could be tackled in a similar way, but by using phonographic recording. Finally, he turned to vowel sounds and would oppose observation with uninformed experiment.

In all these experimentations, false judgment is a recurrent feature. Stumpf explained in the preface to the first volume of *Tonpsychologie*:

The physicist seeks the motives of false judgements only in order to eliminate them. The physiologist as such is perhaps concerned with them for his speculations concerning unknown processes in the brain. To the psychologist, they are essential in that they help him elucidate the coming-about and conditions of judgements as such. In unpractised observers, whom the physiologist rejects from the outset, he studies the influence of practice; and in unmusical people, he studies the conditions of musical feelings. (Stumpf 2020 [1883], Ixii, translation slightly modified).

⁸The German word "Einstellung" means both attitude or mindset, and calibration. The metaphorical transfer between humans and machines does not replicate easily in English. Psychologists and phenomenologists chose "attitude" to translate the word, but their choice does not match technical setups. On attitude and music, see Steege (2021); on "Einstellung" in Stumpf, see Kursell (2021).

⁹See, for instance, Honing (2018).

It was “precisely the curious differences between musical and unmusical natures” (Stumpf 2020, lxi), he added, that supported his research. In other words, false judgments and unmusical subjects provided a key to reducing the complexity of judgment in the realm of music. Experiments with unmusical subjects became a cornerstone in Stumpf’s experimental method. Next to observing how the unmusical judged two tones presented in sequence, he also exposed them to simultaneous tones. This became the main topic in the second volume of *Tonpsychologie* and a key to his influential concept of “fusion” (*Verschmelzung*, see Stumpf (1890)). The method for working on fusion proceeded in two steps. First, subjects were tested about their access to the concept of pitch as explained above. A person who could not answer the question which of two notes was higher would normally be disqualified to judge anything musical. In Stumpf’s setup, however, they qualified for further experimenting, as he needed them for working with two groups: one of them “musical,” the other “unmusical.” Then he exposed both the unmusical and musical subjects to two simultaneously played musical notes. He chose intervals that differed with respect to their consonance and dissonance, speaking in terms of Western tonal harmony.

The question Stumpf asked the subjects in the experiments on tone fusion was not whether they heard a consonance or dissonance, but whether they heard one or several tones. Again, that question is remarkable in how it reduced the complexity of the potential musical background in the subjects’ answers. The answers of “one” or “many” situated the question below a level that already assumed Western tonal music for its framing. Much of the charm of Western music depends on the melting or diverging in simultaneously produced voices—from the choir singing in unison to personalities on stage, like a Marquis de Posa and title hero Don Carlos in Verdi’s 1884 opera singing in parallel thirds and sixths. Those intervals fuse just enough to show two distinct individuals joining in one movement. Music theory was lacking the vocabulary for such features, instead taking the notes on the page as a point of departure: they unambiguously showed whether one or two distinct voices or pitches had to be involved. The category of consonance and dissonance, then, addressed a classification of intervals, rather than their effects in context.

Stumpf’s questions about tones did not depart from music teaching, or, as its elementary level was called in German, *Musiklehre*. Instead, he built on research by Hermann von Helmholtz also when it came to music theory. Helmholtz’s book *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (1863; first English translation 1875) was notorious for providing an explanation of the complementary notions of consonance and dissonance. Music theorists held Helmholtz’s theory of beats that caused roughness in the frequency compounds of simultaneously given tones to favor dissonance, while not explaining the effect of consonance itself. Helmholtz replied in the preface to the third edition that he had never aimed at providing a natural foundation for Western music.¹⁰

¹⁰On Helmholtz’s notion of consonance and dissonance in comparison to Stumpf’s, see Kursell (2008).

Stumpf took another observation in Helmholtz's treatise as his point of departure, namely the amazement regarding the fact that any sound sources can be distinguished at all. Helmholtz used the metaphor of waves on the surface of water to describe the problem. Looking at a water surface in motion, the eye can discriminate directions in the motion and sometimes even discern how many waves intersect in a spot and whence they come. The ear, in contrast, distinguishes only a small spot of such a surface and instead calculates the presence of waves like a mathematician (1954 [1877],¹¹ 36f.).

Interested in the mental operations involved in recognizing tones, Stumpf devised a question posing a simple alternative: do you hear one sound or many? This did not just shift a basic operation of psychophysics to a genuinely psychological task. By asking this question to his two groups of subjects, he also avoided having the subjects—those with and those without a musical background—depend on the vocabulary and knowledge of music for their answer.¹² He then varied the stimuli, always choosing two musical tones but changing their distance or, in other words, the musical interval the tones constituted. The musical intervals, while providing the choice of stimuli, were thus emptied of their musical meaning. Within the system of tonal music (i.e., the music of roughly the year 1600 until Stumpf's own time), the correct answer would always be two tones. All subjects, however, occasionally did not recognize an interval as a manifold. The answers of those subjects for whom musical theory was inaccessible, the unmusical (*Unmusikalische*) in Stumpf's words, in particular shed new light on the reactions from the other group. They tended to hear one sound as the degree of consonance, in musical terms, became higher.

The musically able subjects were, in turn, unable to distinguish the application of their musical knowledge. As a consequence, they could not separate their ability to identify tones as musical notes in certain defined relationships from an immediate sensation. They would, for instance, react to the distinction between consonance and dissonance as it is made in music theory, identifying the two tones accordingly as two consonant or dissonant notes. But they frequently did not identify two notes in the interval of an octave as "many" tones and thought instead they heard just one. Each group could thus be found lacking. The unmusical did not further analyze a multiplicity of tones, they only "sensed" the sound; the musical took the analysis to provide an answer to the question of "one or many," without realizing that they also depended on discerning the multiplicity in a hypothetically prior stage.

From these findings, Stumpf inferred that all subjects sensed two simultaneously given notes as one sound to begin with. The unmusical would remain in the state of that sensation. The musical, in contrast, would analyze the sound in accordance with the rules they had acquired. The immediacy with which the musical subjects reacted to the two notes being consonant or dissonant, in fact, operated as an obstacle for detecting the state of sensation. The analysis happened so fast that

¹¹ The reprint of 1954 is based on the 1885 translation of the German fourth edition of 1877.

¹² On how Ernst Wilhelm Weber construes an alternative question to investigate skin sensation as a means to understand the physiological anatomy of the nerves the skin conceals, and on how this modifies the standard account of psychophysics' history, see Hoffmann (2001, 2005).

they would not notice how they sensed the sound, except when the degree of fusion was exceptionally high. From this, Stumpf construed his notion of fusion, which was later developed into further-reaching phenomenological and Gestalt-theoretical assumptions in his own work, and in that of some of his disciples and colleagues.¹³

As the research on fusion shows, the experiment construed a judgment on sound in terms of simple alternatives operating on two levels. Both the question to be decided on—one or many sounds—and the conditions of judgment—with or without musical ability—were conceived this way. Another aspect is that the subjects could not see the sound sources. They reacted to the sound exclusively, although little was done to shield them from knowledge of a local distribution of the sound, for instance, as would be implied in the sound of specific instruments such as the organ or piano. The logical operation of combining two simple alternatives was at the core of the experiment, and it is this combination that marks a major step in Stumpf's formalization for his method of inquiry. If Stumpf had before collected statements in more informal ways, for instance, through writing letters to friends and colleagues or excerpting literature, he now began to work more systematically with experimental subjects.¹⁴

In 1885 another shift in his work occurred. Attending a performance by non-European musicians in Halle, Stumpf realized that he himself was now in the position of the unknowledgeable listener. The performance appeared to him like some howling and rattling, although he was convinced that this judgment was unjust. He seized the occasion to work with two of the musicians, the singer Nuskilusta and another whose name is not known. These two Nuxalk First Nation singers from British Columbia patiently auditioned with Stumpf in individual sessions. Stumpf did his best to make notes, as he wrote in a paper on this encounter in *Vierteljahrsschrift für Musikwissenschaft* (Stumpf 1886). However, he realized that neither his note paper nor his mind were up to properly marking distinctions relevant to the two singers. A second performance, then, already made a different impression on Stumpf: he meant to hear some singers deviating from what Nuskilusta had taught him. As he remarked tongue in cheek to his readers, unmusical individuals were not a privilege of Western music (Stumpf 1886, 421).

Between the encounter at Halle and the beginning of his research on speech sounds, Stumpf's work on auditory cognition explored the question whether music gave further insights into the mind making sense of it. He eventually founded the Berliner Phonogramm-Archiv, which was to become the largest collection of phonographic wax cylinder recordings worldwide (Ziegler 2006). However, enticing as the prospect might have been to have a multitude of musics to experiment with, all of which followed different implicit rules, recorded sound did not allow

¹³On Stumpf and fusion in the context of phenomenology, see, for instance, Rollinger (1999); on Gestalt psychology in the early years of the Berlin Institute of Psychology, see Ash (1995) and Klotz (1998); on Stumpf's perspective on the interpretations of his notion of fusion, see, e.g., Kaiser-el-Safti (2011).

¹⁴On the research that Stumpf based on collecting statements from acquaintances and colleagues, see Kursell (2019).

experimenting on judgment by comparing groups of initiated listeners with groups who were not. The turn to phonetics eventually brought experimental subjects back into the Institute. As I shall argue below, Stumpf's interest in judgment now also migrated into the material structure he devised for his experimentation. Between his work on music and the study of language sounds, he published several philosophical papers dealing with, among other things, judgment as an epistemological and cognitive problem. With the experimental setup for his phonetics research, then, he practiced a rigorous and controlled way of judging that offered new perspectives on how to provoke judgment for the purpose of empirical scientific investigation.

5.3 A Two-Level Practice of Judging Judgment

The first experiment with unmusical subjects marks the instantiation of what could be called a “practical epistemology,” in Stumpf's own terms. He coined the term in a lecture on logic held at Halle and preserved among the papers of Edmund Husserl. Husserl took notes in 1887 and received a printed version, a so-called “Diktat” (i.e. a text to be dictated), in the following year, 1888 (Fisette 2015a, b; Rollinger 1999, 2015; Schuhmann 1996). While the lecture is considered to lean heavily on those of Stumpf's own teacher Brentano, the term “practical epistemology” is considered to be his own (Rollinger 2015, 77; Schuhmann 1996 on Stumpf's dependence on Brentano more generally). It expresses his opposition to a merely formal approach to logic and asks about uses of logic. Logic is defined in the beginning of the lecture as *Kunstlehre*—to be translated, following Rollinger (2015), as the “instruction to practice an art,” namely the art of correct judgment. The lecture on psychology of the winter semester 1886–87, equally preserved in Husserl's notes, comes back to this understanding of logic:

Logic must go back to the essence of judging, to the different classes of judgments, the expression of them in language, which is indeed also a psychological function. It must also sort out different motives of judging, attend to motives of feeling, habits, exhibit the origin of prejudices, etc. A logic that would abstain from this, a purely formal logic, would otherwise be useless from the outset. (Rollinger 2015, 83 trans. slightly modified)¹⁵

This reads like an outline to Stumpf's work on auditory cognition all the way through, from the first experiments about the reliability of judgment after 1873 up to 1926, when he published his book on speech sounds. Logic, for Stumpf, was not an aim in

¹⁵The German is also given in Rollinger (2015, 83, n. 19): “Die Logik muss zurückgreifen auf das Wesen des Urteilens, auf die verschiedenen Klassen der Urteile, den Ausdruck derselben in der Sprache, die ja auch eine psychologische Funktion ist. Sie muss auch verschiedene Motive des Urteilens auseinanderhalten, die Gefühlsmotive, Gewohnheiten beachten, die Entstehung der Vorurteile aufzeigen etc. Sonst würde eine Logik, die davon absähe, eine rein formale Logik, von vornherein nutzlos sein.” Rollinger emphasizes that Husserl borrowed the notion of “antipsychologism” from Stumpf and not only exempted his mentor from it, but also “allows for logic as *Kunstlehre* to be dependent on psychology and only argues that a pure logic, strictly as a theoretical discipline, must be seen as free of all psychology.” See also Textor (2020).

itself, but as a *Kunstlehre* it had a purpose. The lecture discussed and dismissed other purposes, such as defining logic as either concerned with *thinking*, which Stumpf declares instead to be a matter of psychology, or with *concluding*, which for Stumpf would make it a task of assessing knowledge by means of proof. As the art of practicing correct judgment, logic shared the concern about judgment with psychology. Whether “useful” in the sense of “practicing the art” or useless, Stumpf was critical about the idea that psychology would explain logic or embed it into its own study of the mind. Instead, his interest in judgment overarched two parallel activities: the study of judgment, and the elaboration of methods for doing so. Pushing this further, one could say that Stumpf’s practice of logic included experiment.

In the light of these deliberations, the first volume of *Tonpsychologie* presents a parting of ways. Discussing sensory judgment more generally, Stumpf mused on what psychophysics added to grasping the reliability of judgment:

It would, incidentally, be *a priori* conceivable that yet another constant would have to be added to the specified conditions of the subjective reliability for each individual. [...] If we assume that all previously specified conditions are maximally favourable for a judgement about the equality of two impressions, the question would be whether we would in this case notice every difference, be it ever so small. If not, there would be a threshold that the difference in sensations would have to cross over in order to be discerned as such. This threshold would not have to be dependent on the aforementioned and empirically familiar changeable conditions, but should rather be noted as a peculiarity of the mental (central) organism, as a constant coefficient of discrimination (more generally: of judgement), perhaps variable between individuals. The question, however, can hardly be decided experimentally, for there is, strictly speaking, simply no maximally favourable state for those empirical conditions. They can rather by their nature operate more favourably into infinity. (2020, 21, trans. slightly modified)

Psychophysics was thus caught up in not having and never reaching ideal conditions for experimentation.¹⁶ The quantitative premise that could be tied to what Stumpf identified as its main type of question—same or different—would never be accessible to the ideal conditions it presupposed. More importantly, Stumpf needed a threshold of a different nature. He could not accept pitch to be a homogenous parameter. In *Tonpsychologie* he argued that, at least for those trained in Western tonal music, pitch implied values separated by thresholds beyond which recognition tilted toward one or the other of two neighboring values; it did not imply a fine-grained but compact line between any two values. What is more, the highly developed ability to distinguish pitch in musically trained individuals did not concern the mere question of same or different, but what in music teaching was called “intonation,” that is, the possibility of indicating a value’s closeness to an intended “correct” value. Experimentation that disregarded these features in the musically trained mind was flawed from the outset.¹⁷

Stumpf’s own practice instead proposed to ask what he called “qualitative” questions. Recall the experiment with the unmusical, where he explained, as quoted above, that he “preferred the question ‘which tone is higher?’ to that of ‘equal or

¹⁶The unattainable ideal of experimental conditions has been discussed with the example of Wundt’s experiments on reaction time, in which the subject becomes the main obstacle for assessing the subject’s reactions (Schmidgen 2014).

¹⁷On the debate more generally, see Hui (2013).

different?’,” for he believed this would give him “insight into the general conditions of the qualitative judgement” (2020, 201). The qualitative method could not operate with the core of psychophysics, or a parametrization of sensation in correlation with measured stimuli. The field that interested Stumpf lacked such homogeneity. He described the way in which the realm of pitch was organized for those educated in Western tonal music like a land surveyor’s perspective: nineteenth-century music required a standpoint and, seen from there, the recognition of signposts, rather than a parametrization of pitch.

Stumpf’s own two-level comparison of judgments harmonized with the development of his logic from Brentano’s. From the lectures Stumpf heard from Brentano between 1865 and 1868, he could take a definition of judgment stating that all judgments are “reducible to positive or negative existential judgments” (Schuhmann 1996, 111). Brentano furthermore used a distinction of matter and content, as Karl Schuhmann has explained in a paper on Stumpf as a disciple of Brentano:

The whole complex of presentations underlying the judgment [Brentano] called the judgment’s matter and the act of affirmation or negation he termed the judgment’s form or quality. Further he posited the judgment’s content which he defined as that which is accepted or rejected in the judgment (the immediate target of affirmation or negation, as it were). Such judgmental contents are linguistically expressible in infinitival clauses or in that-clauses. This notion of a content allowed Brentano also to explain so-called indirect judgments of the type ‘it is possible, necessary, true, wrong that ---’ by referring to their content. Thus the judgment ‘it is possible that A exists’ has as its presentational matter A and as its content the possibility of A’s existence. (Schuhmann 1996, 111)

The distinction between preparing the way in which the matters to be judged were phrased, and the ensuing positive or negative judgment, obviously appealed to Stumpf, who was studying law when he heard Brentano lecture for the first time. He later formulated himself the two steps in judging, taking over notions from Bolzano and his other former mentor, Lotze. Stumpf proposed calling the content “Sachverhalt” (i.e., state of affairs). This notion stemmed from German legal practice, where it described the preparation of the file encompassing everything that the judge was entitled to take into consideration for the judgment: “what is not in the file, is not in the world.” This practice implied the separation of two steps in judging. The matter to be judged was first prepared and documents gathered, so as to be presented as the “state of affairs” in the file. The final judgment, then, answered to the state of affairs, not to the matter beyond the confines of the court. In Stumpf’s time, the written file had been replaced by hearing statements before court.¹⁸

This foundation for the distinction between matter and content or state of affairs implies that the question of whether or not a statement is true or false cannot reach any rationalization beyond the content of the judgment. Stumpf could find support for this stance in Brentano. “According to Brentano,” Arkadiusz Chrudzimski writes,

¹⁸The Latin phrase is quoted in Vismann (2011). On the philosophical notion of Sachverhalt as embedded in German legal practice, see Smith (1978). On how Stumpf’s notion of Sachverhalt evolves from gathering statements to experimentation and on how the development of an actual hearing in court may have impacted Stumpf’s interest in acoustics, see Kursell (2022).

“a judgment is not true when it coincides with a part of reality, but when it could also be made by someone who judges based on evidence” (Chrudzimski 2015, 178). This “epistemic notion of truth,” he continues, entailed that Brentano not only could dispose of propositional truth makers, but of any kind of truth makers.¹⁹

Stumpf’s method of studying judgment privileged instances in which considering a judgment’s truth as the basis for further elaboration is irrelevant.²⁰ He took the subjects to be judging as best they could, based on their individual epistemic situatedness or conditions of judgment (*Urteilsbedingungen*). He then compared the outcomes of those judgments by looking at more than one individual. He thereby formed what is in focus for this chapter: groups who share elements of that epistemic situatedness. Rather than defining those elements, however—and this is central to his method’s foundation in Brentano’s logic—the shared element was reduced to being on one side of a yes-or-no alternative.

The members of the group shared that they all did not have some feature defining a second group. That feature could be very simple, such as that group A can tell which of two tones is higher while group B cannot do so; group A does know what sounds are used in the experiment while group B does not; group A is familiar with such and such a regional musical practice while group B is not. As is apparent, statistical relevance was not a defining feature, nor was random choice of the individuals: Stumpf was the only individual in the group of those not familiar with the music of the Nuxalk. The one defining feature was even used for categorizing one group without the second being investigated in a paired setup: the uninformed individuals were not systematically compared to informed subjects. That is to say, the control group, in Stumpf’s case, emerges directly from a logical operation.

The interference device with its distributed architecture gave this logical operation a new, material shape. The judgments Stumpf studied were no longer based on genuinely invalid premises, but rather on arbitrarily induced premises invalidating the judgment. If, for the informed observers, the acoustic topography of language sounds was what they should observe in detail, the uninformed were supposed to resort to everything they were left with in the state of an induced lack. Their notes from the listening task display this function. While the controlled ignorance cut the subjects off from correct judgments of the sounds’ origins, this ignorance not only allowed them to judge the sound in an unbiased way, the comments they added on the notes also made them explicate on which other conditions they fell back for their judgments. Framed by the task handling their controlled ignorance, they added insights on their motives, habits, and prejudices. In short, Stumpf’s method can be summarized as creating situations in which subjects who could or could not judge truthfully were confronted with objects that were prepared for response in a controlled way, rather than a truthful way. What began as his interest in “false”

¹⁹ See also Rollinger (2015, 80), on truth in Stumpf’s logic.

²⁰ For instance, in Stumpf (1906b, 50): “Die Eigenschaft, um derentwillen wir von notwendigen Urteilen im logischen Sinne des Wortes sprechen, ist nicht [die] psychologische, reale Notwendigkeit. Sie ist eine immanente Eigenschaft des Urteils in Hinsicht seines Inhaltes, als des Sachverhaltes. Diesem kommt sie zu, nicht dem Urteilsakt.” See also Stumpf (1906a).

judgment developed into a method investigating judgment based on the distinction between two alternative and mutually exclusive conditions of judgment.

5.4 Conclusion

This chapter has discussed the roles that groups of subjects played in Carl Stumpf's experimental practice or—stretching the etymology somewhat—their roles and counter-roles. Stumpf began working with groups of subjects long before he researched vowels, in his research on judging musical tones. Although those groups do not match the requirements identified by Dehue and spelled out in research on the history of the control group more generally (Dehue 1997, 2001, 2005; Chalmers 2001), and although he did not use the term “control group” even in 1926, by the time the notion gained currency, his research practice has shed new light on it. More specifically, Stumpf transfers one basic feature of control experiments to his psychological investigation: he reduces the claim that “all other factors remain the same” to a simple alternative that he eventually could control arbitrarily. Rather than taking all sentiments and feelings of music-listening into account, he split all music listeners into two groups, according to the criterion of whether they could or could not distinguish the higher of two notes. In this case, his homogeneous Middle-European population easily granted that his subjects would share many features otherwise. The musical and unmusical subjects, for instance, were all eloquent, had access to erudition, were exposed in various ways to music, etc. The uninformed subjects in the vowel experiment were first tested with regard to their general ability to react to the apparatus. Individuals who would not have accepted the transmitted sounds to begin with were thus not admitted to the experiment. In other words, the functioning of the logical operation at the core of this method had to be carefully handled, even though its explicitness varies greatly.

The story this chapter has been telling about the history of experimental psychology diverges considerably from the standard narrative emphasizing psychophysics, and in particular from the telling of Edwin Boring. Stumpf's notion of psychology welcomed experimental methods while rejecting the exclusive methodological choice of correlating sensations to outer stimuli, which had made psychophysics the center of psychology's alleged auto-historiography as instantiated by Boring (e.g., Boring 1929, 1942).²¹ Instead, Stumpf's psychology was based on an immanent approach and his main object of inquiry was judgment. For his practical epistemology, he worked with two parallel strands of developing psychological methodology. In his experimental work, he provoked judgments that he held against the conditions in which they were made. Step-by-step, this method took on a systematic character and culminated in the setup for the vowel experiment, which went so far as to induce

²¹ That Edwin Boring started out from the 1930s electrified psychological laboratory, when he made researchers like Helmholtz and Stumpf predecessors of, e.g., S.S. Stevens, has not yet been considered sufficiently. See, e.g., Boring (1938), and cf. Kursell and Schäfer (2018).

ignorance. This method turned this intervention via induced ignorance into a methodological device that became customary to the notion of control group as it has been discussed here. In parallel, he developed his own notion of logic as a practice. This theoretical backing, despite not proceeding at the same pace, remained constantly connected to his practical work. He anticipated this making him drift away from logic proper, when he wrote in the preface to his book *Die Sprachlaute* (p. v.): “The philosopher who will pick up this book, will shake his head in incomprehension and lay it aside again quickly.”

References

- Ash, Mitchell. 1995. *Gestalt Psychology in German Culture 1890–1967: Holism and the Quest for Objectivity*. Cambridge/New York: Cambridge University Press.
- Boring, Edwin G. 1929. *A History of Experimental Psychology*. New York: Appleton-Century.
- . 1938. Perspective. In *Hearing. Its Psychology and Physiology*, ed. Stanley Smith Stevens and Hallowell Davis, v–vii. New York/London: Wiley/Chapman & Hall.
- . 1942. *Sensation and Perception in the History of Experimental Psychology*. New York: Appleton-Century-Crofts.
- . 1954. The Nature and History of Experimental Control. *The American Journal of Psychology* 67 (4): 573–589.
- Chalmers, Iain. 2001. Comparing Like with Like: Some Historical Milestones in the Evolution of Methods to Create Unbiased Comparison Groups in Therapeutic Experiments. *International Journal of Epidemiology* 30 (5): 1156–1164.
- Chrudzinski, Arkadiusz. 2015. Carl Stumpf über Sachverhalte. In *Philosophy from an Empirical Standpoint: Essays on Carl Stumpf*, ed. Denis Fiset and Riccardo Martinelli, 173–202. Leiden/Boston: Brill Rodopi.
- Dehue, Trudy. 1997. Deception, Efficiency, and Random Groups: Psychology and the Gradual Origination of the Random Group Design. *Isis* 88 (4): 653–673.
- . 2001. Establishing the Experimenting Society: The Historical Origin of Social Experimentation. *The American Journal of Psychology* 114 (2): 283–302.
- . 2005. History of the Control Group. In *Encyclopedia of Statistics in Behavioral Science*, ed. Brian S. Everitt and David C. Howell, 829–836. Chichester: Wiley.
- Deutsches Wörterbuch* von Jacob Grimm und Wilhelm Grimm, digital version at Wörterbuchnetz of Trier Center for Digital Humanities, Version 01/21, <https://www.dwds.de/d/wb-1dwb>. Accessed November 6, 2021.
- Fisette, Denis. 2015a. The Reception and Actuality of Carl Stumpf: An Introduction. In *Philosophy from an Empirical Standpoint: Essays on Carl Stumpf (Studien zur Österreichischen Philosophie, ed. Marco Antonelli, vol. 46)*, ed. Denis Fiset and Riccardo Martinelli, 11–53. Leiden/Boston: Brill and Rodopi.
- . 2015b. Archivalia: Introduction. In *Philosophy from an Empirical Standpoint: Essays on Carl Stumpf (Studien zur Österreichischen Philosophie, ed. Marco Antonelli, vol. 46)*, ed. Denis Fiset and Riccardo Martinelli, 423–432. Leiden and Boston: Brill and Rodopi.
- Helmholtz, Hermann L.F. 1954. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. New York: Dover.
- Hiebert, Erwin. 2014. *The Helmholtz Legacy in Physiological Acoustics*. Cham et al.: Springer.
- Hiebert, Elfrieda, and Erwin Hiebert. 1994. Musical Thought and Practice: Links to Helmholtz’s Tonempfindungen. In *Universalgenie Helmholtz. Rückblick nach 100 Jahren*, ed. Lorenz Krüger, 295–311. Berlin: Akademie Verlag.
- Hoffmann, Christoph. 2001. Haut und Zirkel. Ein Entstehungsherd: Ernst Heinrich Webers Untersuchungen ‘Ueber den Tastsinn’. In *Ansichten der Wissenschaftsgeschichte*, ed. Michael Hagner, 191–223. Frankfurt am Main: Fischer.

- . 2005. *Unter Beobachtung. Naturforschung in der Zeit der Sinnesapparate*. Göttingen: Wallstein.
- Honing, Henkjan, ed. 2018. *The Origins of Musicality*. Cambridge, MA: MIT Press.
- Hui, Alexandra. 2013. *The Psychophysical Ear: Musical Experiments, Experimental Sounds, 1840–1910*. Cambridge, MA: MIT Press.
- Jackson, Myles W. 2006. *Harmonious Triads. Physicists, Musicians, and Instrument Makers in Nineteenth-Century Germany*. Cambridge, Mass., London: MIT Press (= Transformations: Studies in the History of Science and Technology).
- Kaiser-el-Safti, Margret. 2011. Einleitung. In *Carl Stumpf: Erkenntnislehre*, ed. Carl Stumpf Gesellschaft, 5–45. Lengerich: Pabst Science Publisher.
- Klotz, Sebastian, ed. 1998. *Vom tönenden Wirbel menschlichen Tuns. Erich M. von Hornbostel als Gestaltpsychologie, Archivar und Musikwissenschaftler*. Berlin: Milow: Schibri-Verlag.
- Kursell, Julia. 2008. Hermann von Helmholtz und Carl Stumpf über Konsonanz und Dissonanz. *Berichte zur Wissenschaftsgeschichte* 31: 130–143.
- . 2013. Experiments on Tone Color in Music and Acoustics: Helmholtz, Schoenberg, and Klangfarbenmelodie. In *Music, Sound, and the Laboratory From 1750 to 1980*, ed. Alexandra Hui, Julia Kursell, and Myles W. Jackson, vol. 28, Osiris, 191–211. Chicago: University of Chicago Press.
- . 2018. *Epistemologie des Hörens: Helmholtz' physiologische Grundlegung der Musiktheorie*. Paderborn: Fink.
- . 2019. From Tone to Tune: Carl Stumpf and the Violin. *19th Century Music* 43 (2): 121–139.
- . 2021. Coming to Terms with Sound: Carl Stumpf's Discourse on Hearing Music and Language. *History of Humanities* 6: 35–59.
- . 2022. Carl Stumpf and the Phenomenology of Musical Utterances. In *The Oxford Handbook of the Phenomenology of Music Cultures*, ed. Harris M. Berger, Friedlind Riedel, and David VanderHamm, C30.P1–C30.N24. Oxford University Press (online publication).
- Kursell, Julia, and Armin Schäfer. 2018. Elektronische Musik für Radios von John Cage, Karlheinz Stockhausen und Michael Snow. In *Radiophonic Cultures*, ed. Ute Holl, vol. 1, 135–149. Heidelberg: Kehrer.
- Martinelli, Riccardo. 2015. A Philosopher in the Lab. Carl Stumpf on Philosophy and Experimental Sciences. *Philosophia Scientiae* 19 (3): 23–43.
- Pantalony, David. 2005. Rudolph Koenig's Workshop of Sound: Instruments, Theories, and the Debate over Combination Tones. *Annals of Science* 62: 57–82.
- . 2009. *Altered Sensations. Rudolph Koenig's Acoustical Workshop in Nineteenth-Century*. Paris, Dordrecht et al: Springer.
- Rheinberger, Hans-Jörg. 1997. *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Stanford: Stanford University Press.
- . 2023. *Split and Splice: A Phenomenology of Experimentation*. Chicago/London: University of Chicago Press.
- Rollinger, Robin D. 1999. *Husserl's Position in the School of Brentano*. Dordrecht/Boston/London: Kluwer.
- . 2015. Practical Epistemology: Stumpf's Halle Logic (1887). In *Philosophy from an Empirical Standpoint: Essays on Carl Stumpf (Studien zur Österreichischen Philosophie, ed. Marco Antonelli, vol. 46)*, ed. Denis Fiset and Riccardo Martinelli, 75–100. Leiden/Boston: Brill and Rodopi.
- Scherer, Wolfgang. 1989. *Klavier-Spiele. Die Psychotechnik der Klaviere im 18. und 19. Jahrhundert*. München: Wilhelm Fink.
- Schickore, Jutta. 2019. The Structure and Function of Control in the Life Sciences. *Philosophy of Science* 86 (2): 203–218.
- . 2021a. The Place and Significance of Comparative Trials in German Agricultural Writings Around 1800. *Annals of Science* 78 (4): 484–503.
- . 2021b. Methodological Ideas in Past Experimental Inquiry: Rigor Checks around 1800. *Intellectual History Review* 33: 267–286.
- Schmidgen, Henning. 2014. *Hirn und Zeit: Die Geschichte eines Experiments*. Berlin: Matthes und Seitz.

- Schumann, Karl. 1996. Carl Stumpf 1848–1936. In *The School of Franz Brentano*, ed. Liliana Albertazzi, Massimo Libardi, and Roberto Poli, 109–129. Dordrecht: Kluwer.
- Smith, Barry. 1978. Law and Eschatology in Wittgenstein's Early Thought. *Inquiry* 21 (1–4): 425–441.
- Steege, Benjamin. 2012. *Helmholtz and the Modern Listener*. Cambridge: Cambridge University Press.
- . 2021. *An Unnatural Attitude: Phenomenology in Weimar Musical Thought*. Chicago, London: University of Chicago Press.
- Stumpf, Carl. 1873. *Über den psychologischen Ursprung der Raumvorstellung*. Leipzig: Hirzel.
- . 1883. *Tonpsychologie*. Vol. 1. Leipzig: Hirzel.
- . 1886. Lieder der Bellakula-Indianer. *Vierteljahrsschrift für Musikwissenschaft* 2: 405–426.
- . 1890. *Tonpsychologie*. Vol. 2. Leipzig: Hirzel.
- . 1906a. Erscheinungen und psychische Funktionen. In *Abhandlungen der Königlich Preußischen Akademie der Wissenschaften aus dem Jahre 1906*, 40 pp. Abh. IV.
- . 1906b. "Zur Einteilung der Wissenschaften." *Abhandlungen der Königlich Preußischen Akademie der Wissenschaften aus dem Jahre 1906*, Abh. V, 94 pp.
- . 1918. Die Struktur der Vokale. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften* 17: 333–358.
- . 1926. *Die Sprachlaute. Experimentell-phonetische Untersuchungen nebst einem Anhang über Instrumentalklänge*. Berlin: Springer.
- . 2020. *Tone Psychology, vol. 1: The Sensation of Successive Single Tones* (trans: Rollinger, Robin D.). London/New York: Routledge.
- Textor, Mark. 2020. Stumpf between Criticism and Psychologism: Introducing 'Psychologie und Erkenntnistheorie'. *British Journal for the History of Philosophy* 28 (6): 1172–1180.
- Vismann, Cornelia. 2011. Medien der Rechtsprechung. In *and Markus Krajewski*, ed. Alexandra Kemmerer. Frankfurt am Main: S. Fischer.
- Wittje, Roland. 2016. *The Age of Electroacoustics. Transforming Science and Sound 1863–1939*. Cambridge, MA: MIT Press.
- Ziegler, Susanne, ed. 2006. *Die Wachszyylinder des Berliner Phonogramm-Archivs*. Berlin: Ethnologisches Museum, Staatliche Museen zu Berlin – Preußischer Kulturbesitz.

Julia Kursell is Professor of Musicology at the University of Amsterdam. Her research interests include 20th and 21st-century composition, the history of musicology and the relation between music and science. She is co-director of the UvA's Vossius Center for the History of Humanities and Sciences (<http://vossius.uva.nl>).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

