

Experimentelle Psychologie: Psychophysik und Messen in der Psychologie

Ein Lehrbuch mit Selbstversuchen

Von Mark Vollrath

Basierend auf einem Buch von Otto Heller und Hans-Peter Krüger

1. Einleitung

Die Psychologie ist die Wissenschaft vom Erleben und Verhalten der Menschen. Sie ist eine spezielle Art von Wissenschaft, eine Naturwissenschaft. Was ist damit gemeint? Zum einen ein bestimmtes Weltbild – der Mensch ist Teil der Natur und genauso zu erforschen wie Tiere, Pflanzen oder unbelebte Dinge. „Genauso“ bedeutet eine bestimmte Methodik: Man geht empirisch vor, man sammelt also zielgerichtet Daten. Aus den Daten werden Aussagen abgeleitet oder umgekehrt: Aussagen werden mit Hilfe von Daten geprüft. Wissen entsteht also, indem systematisch Daten gesammelt werden.

Diese Beschreibung für Psychologie klingt für viele Studenten abschreckend. Man will doch verstehen, wie der Mensch funktioniert. Man möchte den Menschen verändern können oder ihm helfen, sich selbst zu verändern, sich weiter zu entwickeln. Dieses Verstehen möchte man lernen – nicht das Datensammeln. Keine Sorge, im Bereich der Diagnostik und der Klinischen Psychologie werden Sie über das Verstehen einzelner Menschen eine ganze Menge lernen. In der Allgemeinen Psychologie geht es allerdings weniger um das Individuum, sondern das, was **den** Menschen ausmacht, die grundsätzlichen Eigenschaften und Funktionen. Um das zu verstehen, sollte man Daten sammeln – nur so kann man überprüfbar nachweisen, wie der Mensch funktioniert.

Dies ist am besten an einem Beispiel zu verdeutlichen. Betrachten Sie die folgende Abbildung 1:

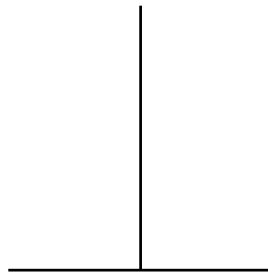


Abbildung 1: Welche der beiden Striche ist länger? Senkrecht oder waagrecht?

Wie man deutlich erkennt, ist der senkrechte Strich länger als der waagerechte. Oder? Nehmen Sie ein Lineal und messen Sie das nach. Interessant: Beide Striche sind tatsächlich gleich lang. (Nebenbei: Sie haben gerade Ihre ersten Daten gesammelt. Und? War gar nicht so schlimm, oder?) Was lernt man aus diesem Vergleich? Die objektiven Eigenschaften von Objekten (Länge von Strichen) scheinen nicht unbedingt mit dem subjektiven Erleben dieser Eigenschaften übereinzustimmen. Oder, etwas allgemeiner, die Wahrnehmung bildet die physikalische Welt nicht einfach nur 1:1 ab, sondern verändert diese. Durch den Vergleich Ihres Urteils („der senkrechte Strich ist länger als der waagerechte“) mit dem Ergebnis der Messung („beide Striche sind gleich lang“) haben Sie eine Erkenntnis gewonnen. Das ist der oben angesprochene eine Teil einer empirischen Wissenschaft: Man sammelt Daten, macht Beobachtungen und versucht, daraus Aussagen abzuleiten, um damit die Wissenschaft weiterzubringen. Durch dieses Vorgehen wird auch das psychologische Wissen gewonnen. Aber hat man damit tatsächlich schon etwas verstanden?

Wir wissen jetzt zwar, dass der senkrechte Strich länger wahrgenommen wird als der waagerechte, aber warum ist das so? Hier kommt der zweite Aspekt einer empirischen Wissenschaft ins Spiel. Man bildet Hypothesen über die möglichen Ursachen und prüft diese mit Hilfe von Beobachtungen und Daten. Zum Beispiel könnte die Überschätzung ja daran liegen, dass der Strich senkrecht ist und für uns als Menschen Größe in der Regel mit Höhe (senkrecht) gleichgesetzt wird. Jemand ist 1.80 Meter groß – in der Senkrechten. Wie kann man die Hypothese prüfen, dass es an der senkrechten Orientierung liegt?

Man könnte die Figur ja um 90 Grad drehen. Und, was passiert? Betrachten Sie Abbildung 2.

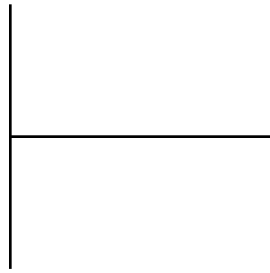


Abbildung 2: Und welcher Strich ist jetzt länger?

Jetzt ist der waagerechte Strich länger. Es kann also nicht an der Ausrichtung liegen. Was haben wir methodisch gemacht? Wir haben unsere Hypothesen mit Hilfe von Daten geprüft. Dazu wurde ein Merkmal, das wir als Ursache vermutet haben, variiert. Der senkrechte Strich wurde durch Drehung zu einem waagerechten Strich gemacht. Alles andere blieb gleich. Wenn es nur an der Ausrichtung gelegen hätte, müsste der Längenunterschied verschwinden. Die Selbstbeobachtung zeigt, dass dies nicht der Fall ist. Also muss die Hypothese verworfen werden. Man nennt dies Falsifikation: der Nachweis, dass etwas (eine Hypothese) falsch ist.

Fällt Ihnen noch eine weitere Hypothese ein, was die Ursache für diesen wahrgenommenen Längenunterschied sein könnte? Denken Sie nach, seien Sie kreativ!

Es könnte sein, dass es etwas mit der Teilung des einen Strichs durch den anderen zu tun hat. Ein langer Strich wird in zwei kurze geteilt und man vergleicht mental die Summe aus zwei kurzen Strichen mit dem einen langen. Wie kann man diese „Teilungshypothese“ prüfen? Schauen Sie Abbildung 3 an:



Abbildung 3: Und welcher Strich ist hier länger?

Jetzt erscheinen die beiden Striche plötzlich gleich lang. Man kann den einen geistig nach oben drehen oder den andere nach unten – gleich lang. Die Teilung scheint also tatsächlich wichtig zu sein. Unsere Hypothese hat sich also bestätigt. Wir wissen damit etwas mehr über die Bedingungen, unter denen dieser Längenunterschied auftritt. Damit ist die Wissenschaft aber natürlich noch nicht zu Ende. Woran liegt es denn, dass der geteilte Strich unterschätzt wird? Hat man als Mensch Schwierigkeiten, Längen zusammenzufügen? Ist es wichtig, dass der Strich genau halbiert wird? Welche Erklärungsmöglichkeiten fallen Ihnen noch ein? Und: wie kann man diese Hypothesen prüfen?

Aber das soll an dieser Stelle nicht weiter vertieft werden – wenn Sie möchten, können Sie das aber gerne in einer eigenen empirischen Studie tun. Das Beispiel sollte vielmehr deutlich machen, wie die Psychologie als empirische Wissenschaft arbeitet:

- Man führt Beobachtungen durch und sammelt Daten, um daraus Hypothesen abzuleiten.
- Man prüft Hypothesen mit Hilfe von Untersuchungen, in denen Bedingungen (mögliche Ursachen) variiert werden und Daten mit Hilfe von systematischen Beobachtungen gewonnen werden, um damit über die Gültigkeit der Hypothesen zu entscheiden.

Eine empirische Wissenschaft ist damit etwas sehr Kreatives – man denkt sich ständig neue Hypothesen aus. Sie hat aber auch den Anspruch auf Wahrheit – man prüft, ob die Hypothesen tatsächlich stimmen. In der Psychologie versteht man auf diese Weise immer besser, wie der Mensch funktioniert.

In dem Kurs sollen Sie beides lernen. Dabei werden wir uns auf den Bereich der menschlichen Wahrnehmung konzentrieren, weil die Wahrnehmung die Basis für alle weiteren Gedanken, Gefühle, Planungen und Handlungen ist. Wie das Beispiel zeigt, weicht das menschliche Erleben systematisch von den physikalischen Eigenschaften ab. Ist dies tatsächlich systematisch? Wie kann man die Abweichung beschreiben? Woran könnte das liegen? Dies sind die inhaltlichen Fragen, die uns in diesem Kurs beschäftigen werden.

Mindestens ebenso wichtig sind die methodischen Fragen: Wie untersucht man die menschliche Wahrnehmung? Wie prüft man Hypothesen schlüssig? Sie werden die Antworten auf diese Fragen bekommen, indem Sie selbst als Forscher arbeiten, mit sich selbst als Versuchsperson. Studiert man nicht auch Psychologie, weil man sich besser selbst verstehen möchte? Dies werden Sie im Lauf dieses Kurses erleben.

Das Vorgehen, um Hypothesen schlüssig zu prüfen, ist das Experiment. Sie sollen im Selbstversuch erlernen, wie ein Experiment funktioniert. Das hat den Vorteil, dass Sie sowohl Versuchsleiter als auch Versuchsperson sind und so beide Sichtweisen kennenlernen können. In Ihrer Abschlussarbeit werden Sie später eigene Versuche planen und als Versuchsleiter durchführen. Um auch dafür erste Erfahrungen zu gewinnen, werden Sie und Ihre Arbeitsgruppe in diesem Kurs auch einen kleinen Versuch mit wenigen Versuchspersonen selbst planen und durchführen. Damit verstehen Sie am Ende des Kurses etwas besser, wie psychologisches Wissen entstanden ist und wissen, wie Sie selbst dazu beitragen können, weiteres Wissen zu erarbeiten. Ist das nicht spannend?

2. Versuch Größeneindruck 1: Weicht das subjektive Erleben von der objektiven Umwelt ab?

2.1. Ein bisschen Theorie

Wenn wir nochmal an das Beispiel aus dem letzten Abschnitt zurückdenken, dann gab es da zwei Aspekte, die unterschieden wurden:

- Der senkrechte Strich, der eine bestimmte messbare Länge hat, und der von der Messung her gleichlang war wie der waagerechte Strich.
- Das Erleben der Strichlänge, wobei uns der senkrechte Strich länger vorkam als der waagerechte.

Dieser Unterschied hat uns zu einer ganz zentralen Frage der Psychologie geführt, nämlich inwieweit das Erleben systematisch von den physikalisch messbaren Eigenschaften der Objekte abweicht. Natürlich möchte man dann auch noch weiter verstehen, welche Konsequenzen das für das Verhalten des Menschen hat.

In der Psychologie hat man eine Zeitlang versucht, das Verhalten direkt als Folge von bestimmten Ereignissen in der Umwelt zu verstehen. Ein ganz einfaches Beispiel: Sie fahren in ihrem Auto an die Uni. Plötzlich rollt ein Ball auf die Straße. Sie machen eine Vollbremsung. Was ist hier psychologisch passiert?

In der Umwelt ist ein Objekt mit bestimmten Eigenschaften aufgetaucht. In der Psychologie bezeichnet man das als „Reiz“ oder „Stimulus“, gerne abgekürzt als „S“. Auf diesen Reiz haben sie mit einer Vollbremsung reagiert, also eine „Reaktion“ (abgekürzt „R“) ausgeführt. Also ganz einfach:

- $S \rightarrow R$

Das Auftauchen eines Reizes führt zu einer bestimmten Reaktion, einem Verhalten. Oder: Immer dann, wenn ein bestimmter Reiz auftaucht, macht der Fahrer eine Vollbremsung.

Es wäre ja vielleicht schön, wenn der Mensch so einfach zu erklären wäre: Man muss nur die richtigen Reize kennen, dann kann man vorhersagen, wie sich Menschen verhalten werden. Ihre eigene Erfahrung zeigt, dass dies nicht so einfach ist. Und vor allem: Man würde ja gerne verstehen wollen, warum der Mensch bei dem Reiz ein entsprechendes Verhalten zeigt. Warum macht man eine Vollbremsung, wenn ein Ball auf die Straße rollt? Die wahrscheinlichste Antwort ist, dass man damit rechnet, dass ein Kind hinter dem Ball herrennen wird. Der Reiz, auf den reagiert wird, ist eigentlich gar nicht der Ball, sondern die Vorstellung oder das Wissen, dass möglicherweise ein Kind folgen wird.

Hier haben wir wieder das Problem, dass der „Reiz“, auf den reagiert wird, nicht dem physikalischen Reiz „Ball“ entspricht, sondern einem bestimmten Wissen, dass mit dem Ball in dieser Situation verbunden ist. Menschen reagieren nicht „einfach so“ auf die physikalischen Eigenschaften von Reizen, sondern der Reiz bedeutet für sie etwas, und darauf reagieren sie. Um diesen Zwischenschritt zu berücksichtigen, hat man das einfache Verhaltensmodell erweitert:

- $S \rightarrow O \rightarrow R$

Ein Stimulus (S) trifft auf einen Organismus (O), der diesen Reiz in einer bestimmten Weise wahrnimmt und dann entsprechend seinem Erleben mit einer Reaktion (R) reagiert. Die etwas eigentümlichen Bezeichnungen kommen daher, dass viele der entsprechenden Untersuchungen mit Tieren durchgeführt wurden, man aber meinte, dass die Ergebnisse auch für Menschen als biologische Wesen (Organismen) gelten würden. Diesen Ansatz werden Sie im Bereich der Allgemeinen Psychologie unter dem Stichwort „Lerntheorien“ noch ausführlich kennenlernen. An dieser Stelle ist nur wichtig, dass hier die im ersten Abschnitt beschriebene Vorstellung

modellhaft dargestellt wird, dass physikalische Reize nicht 1:1 so erlebt werden und auf die Reize direkt reagiert wird, sondern dass der Mensch die Eigenschaften in einer bestimmten Weise erlebt, mit seinem Vorwissen verknüpft und dann entsprechend handelt.

Uns interessiert jetzt, auf welche Weise physikalische Eigenschaften im Menschen abgebildet werden. Einen Eindruck davon sollen Sie jetzt in einem ersten Selbstversuch, dem Experiment „Größeneindruck“, gewinnen. Bitte lesen Sie zunächst die Anleitung und beginnen Sie erst dann mit dem Versuch.

2.2. Anleitung zum Experiment „Größeneindruck“

2.2.1. Vorbemerkungen – ein besonderer Augenblick

Stopp! Denken Sie einen Augenblick nach! Sie werden jetzt etwas ganz Neues erleben! Sie werden kein Lehrbuch mehr lesen, kein Dozent wird etwas erzählen, sondern Sie erfahren jetzt die psychologische Methodik, das Experiment, am eigenen Leib. Das mag jetzt etwas pathetisch klingen, aber es geht auch um nichts Geringeres als die Art und Weise, wie die Psychologie Wissen erarbeitet. Das ist auch der Anspruch einer Universität: Nicht nur Wissen und Fähigkeiten zu vermitteln, sondern die Fähigkeit zum kritischen Denken, zum Generieren von neuem Wissen zu erwerben, darum geht es bei einem Universitätsstudium. Hoffentlich war das auch Ihre Motivation, als Sie sich für ein Studium entschieden haben.

Jetzt zum Versuch selbst: Die Versuchsanweisung oder Instruktion ist der entscheidende Schritt für das Gelingen eines Experiments. Die „Versuchsperson“ oder der „Proband“ soll dazu gebracht werden, bestimmte Dinge auf eine bestimmte Art und Weise zu tun. Hier geht es um ein Experiment, aber ganz ähnlich ist es, wenn Sie als Therapeut eine Veränderung bewirken, in einem Gespräch eine Diagnose stellen oder als Verkehrspsychologe Fahrer davon abhalten wollen, beim Fahren Textbotschaften zu schreiben. Was sagt man und wie sagt man es so, dass sich die Versuchspersonen „richtig“ verhalten? Sie werden im Lauf dieser Übung noch einiges darüber lernen, auf was es bei einer guten Instruktion ankommt. Zwei Informationen sind ganz zentral:

- Was soll man als Versuchsperson machen?
- Wie, auf welche Weise, mit welcher Einstellung, soll man das tun?

Wenn Sie selbst die erste Instruktion schreiben, werden Sie feststellen, dass schon der erste Punkt nicht immer ganz einfach zu erklären ist. Wichtig für den zweiten Punkt ist es, zu erklären, wozu der Versuch dient, was damit erreicht werden soll, denn daraus ergibt sich meistens, wie man dabei vorgehen sollte. Wir werden Ihnen gleich eine Instruktion geben. Achten Sie darauf, ob Sie diese beiden Punkte wiederfinden und wie gut es gelungen ist, Sie zu informieren.

2.2.2. Das Material

Sie brauchen für das Experiment das Versuchsprogramm „Größeneindruck“, das sie im Programm „Experimente zur Übung Allgemeine Psychologie“ (<https://cgi-psych.rz.tu-bs.de>) aus dem Menü auswählen können. Sie brauchen außerdem eine Tabelle (Excel oder OpenOffice), in die Sie ihre Ergebnisse nach dem Experiment übertragen können.

2.2.3. Die Instruktion – worum geht es und was sollen Sie tun?

In dem Experiment wird an einem einfachen Beispiel untersucht, wie physikalische Eigenschaften erlebt werden. Als Eigenschaft wurde dazu die Größe von Geldmünzen gewählt, die Sie aus dem Alltag gut kennen. Sie sollen sich dazu die Geldmünzen vorstellen und dann Ihre Wahrnehmung dieser Größe aus der Erinnerung wiedergeben. Es geht also um Ihr Erleben, nicht um die tatsächliche Größe. Sie sollten jetzt also nicht noch Münzen anschauen, um die Größe möglichst gut wiederzugeben, sondern sich ganz auf Ihre Erfahrungen verlassen. Schließlich soll es dabei ja um Ihr Erleben gehen.

Wenn Sie jetzt gleich die Größe schätzen, sollten Sie das ganz spontan und ohne großes Nachdenken tun. Man könnte sich natürlich Vieles überlegen: „Vielleicht schätzt man kleine Dinge etwas größer ein, große dafür etwas kleiner, also gleiche ich das aus.“ Wenn Sie das so machen, werden Sie ganz andere Ergebnisse bekommen. Und es geht genau nicht darum, wie man über sein Erleben nachdenkt und zu welchen Schlüssen man dabei kommt, sondern um Ihr spontanes Erleben. Auch irgendwelche Hilfen oder Vergleiche (z.B. die Uhr, der Stift, die Tastatur) sollten Sie nicht nutzen. Auch das würde Ihren spontanen Eindruck verfälschen.

Damit Sie Ihr Erleben gut wiedergeben können, ist es wichtig, dass Sie voll auf den Versuch konzentriert sind. Gehen Sie in ein ruhiges Zimmer, in dem Sie alleine sind und nicht gestört werden. Achten Sie auf gute Beleuchtung. Suchen Sie eine entspannte Sitzposition, in der Sie gut arbeiten können. Nehmen Sie sich Zeit, und stellen Sie sicher, dass Sie genug Zeit haben und nicht im Stress sind.

Wenn Sie bei dem Versuch merken, dass Ihre Konzentration nachlässt, dann machen Sie eine Pause. Es geht nicht darum, schnell fertig zu werden, sondern konzentriert und sorgfältig zu arbeiten. Nur dann werden Sie systematische Ergebnisse erhalten.

So, und jetzt zur Aufgabe selbst. Wie sollen Sie Ihren Größeneindruck von einer bestimmten Münze wiedergeben? Das Programm zeigt Ihnen einen Kreis, dessen Größe Sie ändern können. Das hat den Vorteil, dass das Programm direkt messen kann, wie groß ihr vorgestelltes Geldstück ist. Wenn Sie das Mausexplorer von sich weg drehen, wird der Kreis größer. Drehen Sie es zu sich hin, wird der Kreis kleiner.

Sie sollen sich jetzt als Erstes eine 2 Euro Münze vorstellen. Drehen Sie solange am Rad, bis die Größe Ihrer Vorstellung von dem Geldstück entspricht. Machen Sie zur Kontrolle den Kreis etwas größer und etwas kleiner und pendeln Sie sich auf diese Weise auf die für Sie richtige Größe ein. Wenn der Kreis ihrer Vorstellung entspricht, drücken Sie die linke Maustaste. Machen Sie dann weiter mit der nächsten Münze, die in dem Programm angezeigt wird. Das Programm merkt sich die Werte und gibt sie am Ende in einer Tabelle aus, die Sie dann einfach in Excel übertragen können.

Jede Münzgröße wird von Ihnen mehrfach beurteilt werden. Warum? Wenn wir die Werte derselben Münzen vergleichen, bekommen wir einen Eindruck davon, wie genau ihre Empfindung arbeitet. Wenn immer die exakt gleichen Werte herauskommen (im Rahmen der Messgenauigkeit), wären Sie außerordentlich genau. Wahrscheinlich gibt es aber kleine Abweichungen. Wenn wir diese untersuchen, lernen wir etwas über die Genauigkeit der Empfindung. Deshalb wird jede Münze mehrfach vorkommen. Wenn sie merken, dass Sie diese Münze schon hatten, sollten Sie möglichst nicht mit den vorherigen Urteilen vergleichen, sonst kommen Sie ins Nachdenken, was den Versuch stören würde. Versuchen Sie, bei jeder neuen Münze wieder ganz spontan die Größe einzustellen!

Sie erhalten diese Instruktion auch noch einmal vom Programm. Außerdem ist eines ganz wichtig:

- Sie müssen entweder ihre Monitorgröße in Zoll eingeben oder die Monitorhöhe ausmessen in Zentimetern. Nur wenn Sie dies eingeben, kann das Programm die Pixel richtig in Zentimeter umrechnen!
- Fast noch wichtiger: Wenn Sie die Größe eingeben mit Nachkommastellen, müssen Sie einen Punkt verwenden, also z.B. 14.5 cm.

Haben Sie alles verstanden? Hier noch einmal die wichtigsten Informationen im Schnelldurchlauf:

- Es geht um Ihre Vorstellung der Größe von Münzen, ihr subjektives Erleben.
- Achten Sie auf eine ruhige Umgebung und arbeiten Sie konzentriert.
- Urteilen Sie ganz spontan und denken Sie nicht lange darüber nach, was Sie gerade tun.
- Stellen Sie sich jetzt so gut wie möglich eine 2-Euro-Münze vor.
- Schätzen Sie jetzt die Münzgröße, indem sie das Mausrad drehen.
- Pendeln Sie sich auf die richtige Größe ein.
- Drücken Sie dann auf die linke Maustaste. Das Programm merkt sich den Wert.
- Danach wird Ihnen die nächste Münze vorgegeben. Wiederholen Sie den Versuch so oft, wie es das Programm vorgibt.
- Arbeiten Sie konzentriert und sorgfältig und verlassen Sie sich ganz auf Ihre Vorstellung!
- Nach dem Versuch erscheint eine Tabelle mit den Daten. Markieren und kopieren Sie diese und fügen Sie die Daten in eine neue Tabelle in Excel ein. Speichern Sie diese Tabelle!

Und hier geht es zum Versuch:

- <https://cgi-psych.rz.tu-bs.de>

Lesen Sie nach dem Versuch den nächsten Abschnitt und werten Sie dann Ihre Daten aus.

2.3. Ein kleiner Exkurs zur Statistik

Sie haben ein Experiment durchgeführt, um zu untersuchen, wie die physikalische Größe subjektiv erlebt wird. Was ist das Ergebnis dieses Versuchs? Entsprechend der Fragestellung müssen dazu die physikalische Größe und die erlebte Größe verglichen werden. Die physikalische Größe der Münzen kann man in Millimetern ausmessen. Die von Ihnen erlebte Größe wurde ebenfalls in Millimetern gemessen. Diese Daten müssen nur zusammengefasst und dargestellt werden, um das Ergebnis zu beschreiben. Dazu dient die deskriptive Statistik. Bevor wir Ihre Daten auswerten, sollen ganz kurz einige zentrale Berechnungsmethoden eingeführt werden, die wir für die Auswertung benötigen.

2.3.1. Der Mittelwert

Der wohl wichtigste Kennwert, um Daten zu beschreiben ist der Mittelwert. Hier die Formel:

$$\text{Mittelwert} = \frac{\text{Summe der Messwerte}}{\text{Anzahl der Messwerte}}$$

Was sagt der Mittelwert aus? Sie haben jetzt gerade mehrfach die Größe der 2-Euro-Münze eingestellt. Die Messwerte sind nicht völlig gleich (woher das kommt, dazu später). Wenn wir jetzt beschreiben wollen, wie groß für Sie typischerweise die 2-Euro-Münze ist, bilden wir den

Mittelwert. Also: Immer wenn man viele Messwerte zusammenfassen möchte und den typischen Wert wissen möchte, berechnet man den Mittelwert. Fast alle statistischen Aussagen, die man liest, beziehen sich auf Mittelwerte: „Dieser Juni war kälter als im Jahr vorher“ bedeutet, dass der Mittelwert der Temperaturen der Tage im Juni in diesem Jahr kleiner ist als der Mittelwert der Temperaturen der Tage im Juni des Jahres vorher.

Bei diesem Zusammenfassen der Werte gehen natürlich auch Informationen verloren. Dazu ein kleines Beispiel. Öffnen Sie dazu Excel oder ein anderes Tabellenprogramm und geben Sie in zwei Spalten die Werte aus Tabelle 1 ein:

Tabelle 1: Zwei Messwertreihen als Beispiel.

	Reihe 1	Reihe 2
1	1	-20
2	2	-5
3	3	5
4	4	15
5	5	20

Die beiden Messwertreihen sehen sehr unterschiedlich aus. Die erste Reihe hat nur positive Werte, die zweite auch einige negative. Die erste Reihe hat sehr ähnliche, kleine Werte, die zweite sehr unterschiedliche, auch sehr große Werte.

Berechnen Sie jetzt für beide Reihen den Mittelwert. In Excel geht das sehr einfach. Gehen Sie in die Zelle unter der Reihe 1 und tippen Sie die Zeichen in den folgenden Anführungszeichen ein: „=Summe(, – also das „=-“-Zeichen, danach schreiben Sie „Summe“ und dann die offene Klammer „(,“. Markieren Sie jetzt mit der Maus die Messwerte der linken Reihe (linke Taste drücken und gedrückt halten, bis Sie alle markiert haben). Danach geben Sie „)“ ein und drücken Enter. In der Zelle sollte jetzt eine „15“ stehen, die Summe. Wenn Sie jetzt nach oben über die Tabelle sehen, finden Sie dort in einem Feld die Formel, die Sie gerade eingegeben haben (s. die folgende Abbildung). „Summe“ ist der Name der Funktion, die berechnet wird. „C4:C8“ gibt an, welche Werte einbezogen werden. Bei mir standen die Werte in der Spalte C, dort in den Zeilen 4-8.

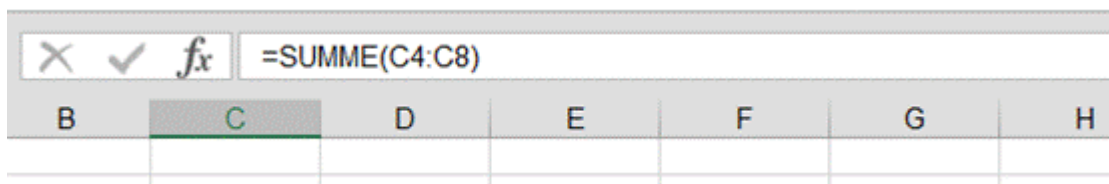


Abbildung 4: Screenshot aus Excel. Oben können Sie die Formel sehen, die eingegeben wurde.

Wiederholen Sie dies für die zweite Reihe. Statt die Formel als Text einzugeben, können Sie auch einfach die Zelle der ersten Reihe kopieren (STRG-C) und in derselben Zeile unter Reihe 2 wieder einfügen (STRG-V). Excel passt automatisch die Formel an, so dass Sie jetzt mit den Werten der zweiten Reihe rechnen. Sie können das im Formelfeld überprüfen.

Wenn Sie alles richtig gemacht haben, sollte in beiden Zellen eine „15“ als Summe stehen. Daraus können Sie jetzt einfach den Mittelwert berechnen, indem Sie diese Summe durch 5, die Anzahl der Messwerte, teilen. In beiden Fällen ist der Mittelwert also 3, und dass, obwohl die Reihen auf den ersten Blick so unterschiedlich sind. Im Mittel sind sie gleich, aber die Ausprägung oder Verteilung der Messwerte ist doch sehr unterschiedlich.

Noch eine kleine Anmerkung: In Excel gibt es sehr viele Funktionen, mit denen man sehr einfach Berechnungen durchführen kann. Es gibt auch die Funktion „Mittelwert“. Gehen Sie in die Zeile unter der Summe, die Sie berechnet haben, und geben Sie dort diese Funktion genauso ein wie die Summe („=Mittelwert(„, Zellen markieren, „)“, Enter). Wenn Sie alles richtig gemacht haben, erscheint jetzt eine „3“. Wiederholen Sie das für die zweite Reihe.

2.3.2. Summe der Abweichungen und die mittlere Abweichung

Das Beispiel zeigt, dass der Mittelwert nur einen bestimmten Aspekt von Daten beschreibt, die typische oder mittlere Ausprägung. Die Messwerte, die zu einem bestimmten Mittelwert führen, können aber sehr unterschiedlich sein. Was bedeutet das? Die Messwerte können alle sehr eng um den Mittelwert liegen, diesem also sehr ähnlich sein, oder auch ganz unterschiedlich sein, aber so, dass im Mittel derselbe Wert herauskommt wie bei der ersten Reihe. Dies kann man sehr gut veranschaulichen, indem man von jedem Wert einer Reihe seinen Mittelwert abzieht. Tabelle 2 zeigt das Ergebnis.

Tabelle 2: Zwei Messwertreihen, die für die Daten aus Tabelle 1 die Differenzen zum Mittelwert zeigen.

Differenz 1	Differenz 2
-2	-23
-1	-8
0	2
1	12
2	17

Sie sollten das selbst in Excel eingeben oder berechnen, um es besser zu verstehen. Gehen Sie dazu in eine freie Spalte in der derselben Zeile, in der ihr erster Messwert der ersten Reihe steht („1“). Sie können hier die Abweichung vom Mittelwert auch als Funktion berechnen, indem Sie folgendes eintippen: „-“ (Markieren Sie jetzt den Messwert) „-“, (Markieren Sie jetzt den Mittelwert) „Enter“. Wenn Sie jetzt nochmal mit dem Cursor in die Zelle gehen, sehen Sie die Formel, die Sie eingegeben haben. Bei mir steht dort z.B. „=C4-C12“, weil der Messwert in der Zelle C4 und der Mittelwert in der Zelle C12 steht. Sie können das jetzt für jeden Messwert der beiden Reihen wiederholen und so alle Werte berechnen. Sie können aber auch die Kopierfunktion benutzen. Allerdings müssen sie dazu vorher die Formel ein kleines bisschen verändern. Gehen Sie dazu in das Formelfeld oben und schreiben Sie vor die Zahl der Zeile des Mittelwerts ein „\$“ (bei mir heißt die Formel dann „=C4-C\$12“. Dieses Zeichen bedeutet, dass beim Kopieren dieser Funktion die Zahl nach dem „\$“ nicht angepasst wird. Das ist wichtig, weil ja der Mittelwert in einer bestimmten Zeile (bei mir 12) steht und bei der Berechnung auch immer auf Zeile 12 zugegriffen werden soll.

Jetzt können Sie die Formel kopieren (in die Zelle gehen, STRG-C). Markieren Sie jetzt die vier Zeilen unter ihrer berechneten Differenz und fügen Sie die Formeln ein (STRG-V). Sie können das jetzt für die zweite Reihe wiederholen, indem Sie einfach in der Spalte neben ihren ersten 5 Differenzwerten die Zellen markieren und die Formel erneut einfügen.

Soweit zur Berechnung, zurück zu den Werten. Man sieht jetzt noch deutlicher als bei den ursprünglichen Messwerten, dass sich in der Reihe 2 die Werte sehr stark in positiver und negativer Richtung von dem Mittelwert unterscheiden. Man kann dies auch grafisch sehr gut veranschaulichen (s. Abbildung 5).

Man möchte diese Eigenschaft der Daten natürlich auch durch einen Kennwert beschreiben. Dieser soll aussagen, wie groß die Abweichungen typischerweise vom Mittelwert sind. Die

einfachste Idee wäre es, einfach den Mittelwert dieser Abweichungen zu berechnen. Mit Hilfe der Formel in Excel geht dies sehr einfach. Allerdings ist das Ergebnis enttäuschend: In beiden Fällen kommt „0“ raus. Das liegt an der Definition des Mittelwerts. Er liegt in der Mitte der Daten, so dass die Abweichungen positiv und negativ sind. Wenn man die einfach aufsummiert und durch die Anzahl der Werte teilt, kommt eben deshalb genau „0“ raus. Große positive und große negative Werte müssten also beide als „groß“ gewertet werden. Das erreicht man, indem man das Vorzeichen weglässt oder die Absolutwerte nimmt. Die Formel dazu in Excel heißt „Abs“ (Sie können also zwei neue Spalten erstellen und in die Zellen jeweils diese Formel eintippen, also z.B. „=Abs(F4)“). Die Formel „Abs“ macht also einfach die Zahlen alle positiv. Das Ergebnis zeigt die folgende Tabelle 3.

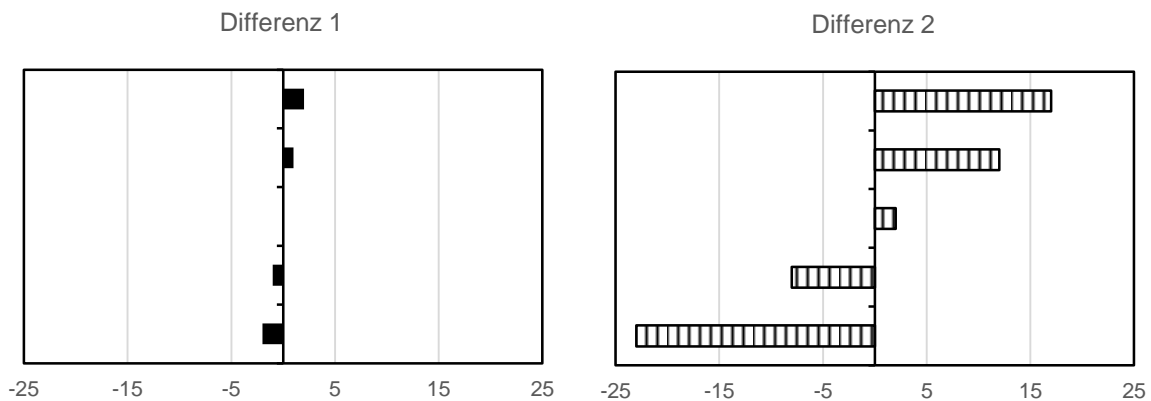


Abbildung 5: Grafische Darstellung der beiden Messwertreihen (Differenzen zum jeweiligen Mittelwert). Die Differenzen wurden der Größe nach geordnet.

Tabelle 3: Absolutwerte der Differenzen zum Mittelwert für die beiden Messwertreihen aus Tabelle 1.

	Absolut 1	Absolut 2
	2	23
	1	8
	0	2
	1	12
	2	17
D	6	62

Wenn man jetzt in dieser Tabelle die Summen der beiden Reihen bildet, ergibt sich in der ersten Spalte („Absolut 1“) 6 und in der zweiten Spalte („Absolut 2“) 62. Diese Summe wird als „D“ bezeichnet (für Summe der absoluten **D**ifferenz). Man erkennt hier sehr gut, dass die Messwerte der ersten Reihe dem Mittelwert relativ ähnlich sind und insgesamt nur um 6 von ihm abweichen, während die der zweiten Reihe insgesamt 62 abweichen. Die Summe D ist also ein einfaches Maß für die Unterschiedlichkeit der Messwerte.

Sie hat einen Nachteil: Die Summe wird immer größer, je mehr Messwerte vorhanden sind. Dies ist nicht schlimm, wenn man zwei Reihen vergleicht, die dieselbe Anzahl von Messwerten haben wie in unserem Beispiel. Häufig liegen aber unterschiedlich viele Messwerte vor. Der Ausweg ist, nicht die Summe, sondern den Mittelwert der absoluten Abweichungen zu berechnen, also die Summe einfach durch die Anzahl der Messwerte zu teilen. Berechnen Sie dies in

Excel für unser Beispiel! Es ergibt sich „1.2“ und „12.4“. Diese mittlere absolute Abweichung nennt man auch „a“. Als Formel:

$$a = \frac{D}{n}$$

Dabei ist D die Summe der absoluten Abweichungen und n die Anzahl der Messwerte.

2.3.3. Der Variationskoeffizient

Die mittlere absolute Abweichung gibt an, wie stark die Messwerte von ihrem Mittelwert abweichen. Nun können die Mittelwerte aber selbst schon ganz unterschiedlich groß sein. Wenn man z.B. vergleichen möchte, ob die Abweichungen bei Schätzungen der Größe einer 2-Euro-Münze anders sind als die bei Schätzungen der Größe eines Fußballs, könnte es Sinn machen, die Abweichung auf die Grundgröße der Objekte, die beurteilt werden, zu beziehen. Das macht der Variationskoeffizient:

$$a\% = \frac{a}{M} * 100$$

Dabei ist a die mittlere absolute Abweichung und M der Mittelwert. Dieser Bruch wird noch mit 100 multipliziert, so dass man a% als „prozentuale Abweichung vom Mittelwert“ interpretieren kann. In unserem Beispiel war in der ersten Reihe a = 1.2 und der Mittelwert 3. Damit ergibt sich für a% (auch das kann man sehr gut in Excel als Formel eingeben):

$$a\% = \frac{1.2}{3} * 100 = 40\%$$

Die Messwerte der ersten Reihe weichen also im Mittel um 40% vom Mittelwert ab. Bei der zweiten Reihe ergibt sich ein a% von 413%. Durch diese Gegenüberstellung wird noch einmal besser deutlich, wie groß die Unterschiede zwischen den beiden Reihen in der Abweichung oder Streuung sind.

2.3.4. Zusammenfassung

Sie haben jetzt folgende statistische Kennwerte gelernt:

Der Mittelwert (M) oder typische Wert der Daten ergibt sich als Summe (S) der Messwerte geteilt durch deren Anzahl (n), also:

$$M = \frac{S}{n}$$

Die mittlere absolute Abweichung (a) gibt an, wie weit die Messwerte um den Mittelwert streuen oder wie unterschiedlich sie sind und wird berechnet als Summe der absoluten Differenzen (D) geteilt durch die Anzahl der Messwerte (n):

$$a = \frac{D}{n}$$

Und schließlich kann die mittlere absolute Abweichung (a) noch auf den Mittelwert (M) bezogen werden, um so die prozentuale Abweichung (a%) zu ergeben (auch als Variationskoeffizient bezeichnet):

$$a\% = \frac{a}{M} * 100$$

Damit haben Sie die wichtigsten Kennwerte kennengelernt und können jetzt den Versuch auswerten.

2.4. Auswertung des Versuchs

2.4.1. Nutzung von Tabellen

Sie haben bereits bei der Berechnung von Kennwerten Excel kennengelernt. Für die Berechnungen und Grafiken werden wir dieses Programm weiter nutzen. In Excel werden die Spalten mit Buchstaben bezeichnet, die Zeilen mit Zahlen. Damit kann man sehr einfach eine bestimmte Zelle auswählen. Die Zelle A3 ist die erste Spalte, 3. Zeile, die Zelle B10 die zweite Spalte mit der 10. Zeile. Wenn Sie Formeln eingeben, können Sie auf diese Weise auch per Hand auf bestimmte Werte verweisen. Das werden wir für die folgenden Berechnungen nutzen.

2.4.2. Berechnung von Mittelwert und Streuung in Ihren Daten

Wenn Sie den Versuch durchgeführt haben, werden Ihre Ergebnisse als Tabelle angezeigt. Abbildung 6 zeigt ein Beispiel. Sie markieren jetzt diese Tabelle mit der Maus und kopieren diese dann (STRG-C; oder rechte Maustaste und „Kopieren“). Dann gehen Sie in die Exceltabelle mit den Auswertungen und dort zum Reiter „Groesseneindruck 1“ (unten).

Dies sind deine Ergebnisse!

Münzen	1 Cent	10 Cent	50 Cent	2 Euro
1. Durchgang	1.39	1.59	2.38	3.17
2. Durchgang	1.19	1.69	2.58	2.78
3. Durchgang	1.29	1.59	2.68	2.97
4. Durchgang	1.39	1.59	2.58	2.68

Abbildung 6: Beispiel für die Ergebnistabelle im Programm.

Sie sehen dort oben links die Tabelle, mit der Sie weiter rechnen sollen (s. Abbildung 7). Markieren Sie diese Tabelle, klicken die rechte Maustaste und wählen „Inhalte einfügen“. Dort wählen Sie aus „Text“. Dann sollte Ihre Tabelle so aussehen wie in Abbildung 8.

Ab hier selbst rechnen! Im Buch ist die Anleitung.

Münzen	1 Cent	10 Cent	50 Cent	2 Euro
1. Durchgang				
2. Durchgang				
3. Durchgang				
4. Durchgang				

Abbildung 7: Tabelle aus Excel, in der Sie weiter rechnen sollen.

Münzen	1 Cent	10 Cent	50 Cent	2 Euro
1. Durchgang	1.39	1.59	2.38	3.17
2. Durchgang	1.19	1.69	2.58	2.78
3. Durchgang	1.29	1.59	2.68	2.97
4. Durchgang	1.39	1.59	2.58	2.68

Abbildung 8: Beispiel für die Tabelle mit Ihren Werten.

Was interessiert uns jetzt an diesen Daten? Was haben wir bei dem Versuch variiert? Das Wichtigste ist sicherlich die Münzgröße. Die Tabelle wurde vom Programm schon nach den Münzen sortiert, sodass in jeder Spalte alle Durchgänge von einer Münzgröße zu finden sind.

Sehen Sie sich jetzt die Werte in Ihrer Tabelle an. Wenn Sie beim Eintragen keinen Fehler gemacht haben und sorgfältig gearbeitet haben, sollten sich die Werte in den vier Spalten deutlich voneinander unterscheiden und die Werte in einer Spalte relativ ähnlich sein. Wie stark unterscheiden sich die Spalten voneinander? Wie groß wird typischerweise ein 1 Cent-Stück von Ihnen geschätzt? Wie genau, zuverlässig, wiederholbar können Sie die Größe schätzen, wie stark variieren Ihre Urteile bei einer Münzgröße also?

Diese Fragen können Sie beantworten, indem Sie die Kennwerte berechnen, die wir im letzten Kapitel kennengelernt haben. Erweitern Sie dazu die Tabelle 3 um die in Tabelle 4 dargestellten Zeilen. Berechnen Sie zunächst die Summe der Werte (z.B. „=Summe(C6:C9)“), danach den Mittelwert. Für den Mittelwert können Sie entweder mit einer Formel die Summe durch 4 teilen oder die Formel „=Mittelwert(C21:C24)“ verwenden. Sie können diese Formeln in die anderen Spalten kopieren. Excel passt dabei automatisch die Spalten an beim Kopieren.

Tabelle 4: Berechnung der statistischen Kennwerte für das Beispiel

Münzen	1 Cent	10 Cent	50 Cent	2 Euro
1. Durchgang	1.39	1.59	2.38	3.17
2. Durchgang	1.19	1.69	2.58	2.78
3. Durchgang	1.29	1.59	2.68	2.97
4. Durchgang	1.39	1.59	2.58	2.68
Summe der Werte	5.26	6.46	10.22	11.60
Mittelwert	1.32	1.62	2.56	2.90
Absolute Abweichung vom Mittelwert	1 Cent	10 Cent	50 Cent	2 Euro
1. Durchgang	0.08	0.02	0.18	0.27
2. Durchgang	0.13	0.08	0.02	0.12
3. Durchgang	0.02	0.02	0.13	0.07
4. Durchgang	0.08	0.02	0.02	0.22
Summe der Abweichungen	0.30	0.15	0.35	0.68
Mittlere Abweichung	0.08	0.04	0.09	0.17
Variationskoeffizient	5.70	2.32	3.42	5.86

Mit den Mittelwerten können wir bereits beschreiben, wie die Münzgrößen eingeschätzt werden. Für die Beurteilung der Genauigkeit der Schätzung brauchen wir die Abweichungen von den Werten, die man wiederum sehr einfach in der Tabelle errechnen kann. Dies sehen Sie im

mittleren Teil von Tabelle 4. Von jedem Originalwert wird der entsprechende Mittelwert abgezogen und das Vorzeichen ignoriert (als Funktion: „=abs([Zelle Einzelwert]- [Zelle Mittelwert])“). Von dieser Tabelle können Sie dann wieder die Summe der Abweichungen berechnen und die mittlere Abweichung. Wenn man die mittlere Abweichung durch den Mittelwert teilt und mit 100 multipliziert, ergibt sich der Variationskoeffizient.

Damit haben wir die wichtigen Kennwerte berechnet für die Beantwortung der oben genannten Versuchsfragen. Damit können wir im nächsten Abschnitt eine erste Antwort geben auf die Frage: Wie gut ist die Schätzung von Münzgrößen?

2.4.3. Antwort auf die Versuchsfragen I: Wie gut ist die Schätzung von Münzgrößen?

Wie gut, wie genau können jetzt die Münzgrößen eingeschätzt werden? Eine erste Antwort auf diese Frage ergibt sich, wenn man den von Ihnen berechneten Mittelwert mit den wahren Größen der Münzen vergleicht. Dazu müssen Sie noch die Abweichungen im unteren Teil der Tabelle in Excel berechnen. Dies ist als Beispiel in Tabelle 5 dargestellt. Hier sind schon die wahren Größen der Münzen eingetragen. In der zweiten Zeile ist die Abweichung (eingestellte Größe minus wahre Größe) dargestellt. Achten Sie hier darauf, dass Sie tatsächlich von ihrer eingestellten Größe die wahre Größe abziehen. Negative Werte bedeuten dann, dass Sie die Größe kleiner eingeschätzt haben, positive Werte, dass Sie die Größe zu groß geschätzt haben. Dies müssen Sie noch über eine Formel berechnen. Dann wird der dritten Zeile diese Abweichung von der wahren Größe in Prozent berechnet ($100 * \text{Abweichung} / \text{wahre Größe}$).

Tabelle 5: Wahre Größe der Münzen in mm und die Abweichung der individuellen Mittelwerte von diesen wahren Größen. Dargestellt ist außerdem die Abweichung in Prozent.

	1 Cent	10 Cent	50 Cent	2 Euro
Wahre Größe	1.63	1.98	2.43	2.58
Abweichung	-0.31	-0.36	0.13	0.33
Abweichung %	-19.08	-18.23	5.36	12.62

Man kann nun diese Zahlen interpretieren. Bei den zentralen Ergebnissen bietet es sich aber auch an, diese als Grafik darzustellen, da man auf diese Weise einfacher direkt sehen kann, wie das Ergebnis zu beschreiben ist. Abbildung 9 zeigt eine Möglichkeit der Darstellung. Hier ist für die verschiedenen Münzgrößen die Abweichung in Prozent dargestellt. Man sieht, dass bei den beiden kleinen Münzen (1 und 10 Cent) die Abweichung relativ groß ist und im negativen Bereich um -20% liegt. Bei den beiden größeren Münzen (50 Cent und 2 Euro) ist die Abweichung dagegen etwas kleiner und liegt im positiven Bereich. Das Ergebnis für das Beispiel würde man damit so zusammenfassen:

- Kleine Münzen werden in ihrer Größe stark unterschätzt, große Münzen etwas überschätzt.

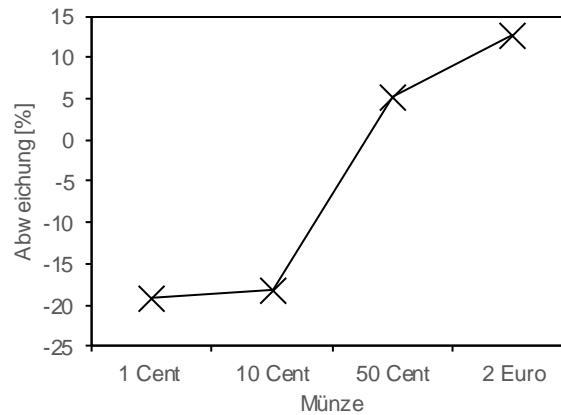


Abbildung 9: Die Abweichung in Prozent in Abhängigkeit von der Art der Münze für das oben dargestellte Beispiel.

Dieses Ergebnis beruht jetzt nur auf einer Versuchsperson. Abbildung 10 zeigt das Ergebnis einer Gruppe von 12 Versuchspersonen, deren mittlere Abweichungen wiederum über die Versuchspersonen hinweg gemittelt wurden. Das Ergebnis stimmt zunächst recht gut mit dem einzelnen Ergebnis aus Abbildung 9 überein. Allerdings zeigt sich in der Gruppe noch eine etwas stärkere Abhängigkeit von der Größe der Münzen:

- Je kleiner eine Münze ist, desto stärker wird ihre Größe unterschätzt. Je größer eine Münze ist, desto kleiner wird die Abweichung von der realen Größe, desto genauer wird also die Größe geschätzt.

Auch Ihr Ergebnis wird mit großer Wahrscheinlichkeit nicht dem der Gruppe entsprechen. Es sollte aber prinzipiell ein ähnlicher Effekt zu erkennen sein. Wenn nicht, woran könnte es gelegen haben? Überlegen Sie, wie Sie den Versuch durchgeführt haben, was Sie sich dabei gedacht haben, ob Sie wirklich spontan und ohne sich groß Gedanken gemacht zu haben, den Versuch durchgeführt haben.

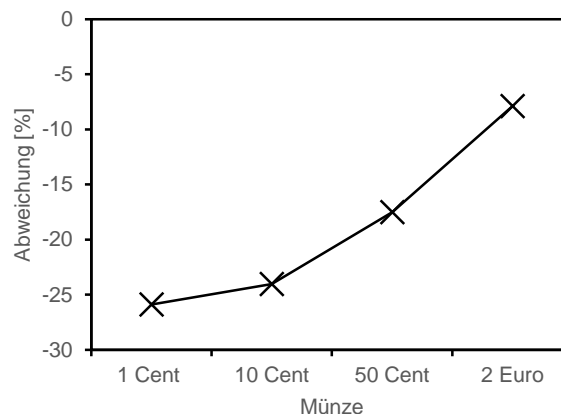


Abbildung 10: Ergebnis eines Versuchs mit $n = 12$ Versuchspersonen. Dargestellt ist für jede Münzgröße der Mittelwert der prozentualen Abweichungen der 12 Versuchspersonen.

Ein typischer „Fehler“ beim Versuch wäre, sich die Münzen auf den Tisch zu legen und beim Einstellen immer direkt mit den Münzen zu vergleichen. Wenn man das tut, wird die Kurve vermutlich waagrecht und sehr nahe an der „0“ verlaufen. Sie hätten dann nicht die Größe aus dem Gedächtnis geschätzt, sondern eine Größe, die Sie sehen, reproduziert, wobei Sie ständig die beiden Größen direkt visuell vergleichen können. Bei diesem Versuch stünde die Genauigkeit der visuellen Wahrnehmung im Vordergrund, nicht aber Ihre im Gedächtnis gespeicherte Vorstellung.

Ein anderer Fehler wäre es, wenn Sie vor dem Versuch schon eine Hypothese gehabt hätten, dass vielleicht kleine Münzen zu klein eingeschätzt werden und dann bei der Durchführung jeweils versucht hätten, diese Fehleinschätzung zu korrigieren. Auch dann würde man eine waagerechte Kurve erwarten. Sie hätten in diesem Fall ihre Vorstellung im Gedächtnis durch Nachdenken korrigiert und so nicht mehr ihr spontanes Urteil genutzt.

Beide Fehler bezeichnet man als „Objektirrtum“. Das, worum es geht, wird in der Psychologie auch als „Objekt“ oder Untersuchungsobjekt bezeichnet. Wir wollten die Vorstellung von Größe im Gedächtnis untersuchen. So war der Versuch aufgebaut und dies war in der Instruktion auch versucht worden, der Versuchsperson ganz klar zu machen. Bei den beiden Fehlern wurde nun das Untersuchungsobjekt verändert. Beim ersten Fehler wurde nicht die Vorstellung im Gedächtnis, sondern die visuelle Wahrnehmung untersucht. Beim zweiten Fehler wurde geprüft, wie man durch Wissen um einen Effekt diesem bewusst entgegenarbeiten kann.

Beide Fehler führen dazu, dass die individuellen Kurven völlig anders aussehen und das Ergebnis dieser Versuchspersonen wertlos ist, da nicht das untersucht wurde, was man untersuchen wollte. Wenn dies bei sehr vielen Versuchspersonen passiert, wird man kein eindeutiges Ergebnis mehr finden können. Deshalb ist es so wichtig, dass alle Versuchspersonen auf die gleiche Weise, ausführlich und verständlich instruiert werden und dass es einem mit der Instruktion gelingt, dass die Probanden verstehen, worum es in dem Versuch geht und wie Sie diesen durchführen sollen.

Bei dieser Art von Versuchen kann es dagegen immer passieren, dass man sich bei einem Reiz vertut, kurz unaufmerksam oder abgelenkt ist und dann ein einzelner Wert völlig anders ist als die anderen Messwerte in dieser Bedingung. Man kann dies bei der Berechnung der Abweichungen vom Mittelwert sehr gut erkennen. Einzelne solche Ausreißer können sich sehr stark auf den Mittelwert auswirken und damit den Kurvenverlauf sehr stark ändern. Bei solchen Ausreißern handelt es sich tatsächlich um unsystematische Fehler. Da dies immer mal passieren kann, wiederholt man die Messungen gerne. Das war auch ein Grund, warum jede Münze viermal eingeschätzt werden sollte. Wenn man das häufig genug tut, sollten sich unsystematische Fehler gegenseitig aufheben, da man dabei vermutlich manchmal unterschätzt, ein anderes Mal überschätzt. Wenn die einzelnen Ausreißer sehr extrem sind, kann es auch sinnvoll sein, sie zu eliminieren, so dass dieser Fehler nicht das Gesamtergebnis bestimmt.

Dabei ist aber immer darauf zu achten, dass man auf diese Weise nicht das Ergebnis verändert. Natürlich darf man nicht die Werte eliminieren, die den eigenen Erwartungen widersprechen. Oder solange Werte eliminieren, bis ein interessantes Ergebnis entsteht. Der Anspruch einer empirischen Wissenschaft ist ja, dass sie ihre Aussagen mit Daten begründet. Daten müssen daher auch ernst genommen werden. Das Eliminieren von Ausreißern muss daher immer gut begründet, nachvollziehbar beschrieben und wirklich nur wenn notwendig durchgeführt werden.

Doch zurück zur Fragestellung: Wie gut ist die Schätzung von Münzgrößen? Die erste Auswertung in diesem Abschnitt hat gezeigt, dass es eine Diskrepanz gibt zwischen objektiver und subjektiver Münzgröße. Diese scheint von der Münzgröße abzuhängen: Je größer die Münze ist, desto eher entspricht die subjektive Schätzung der objektiven Schätzung. Bei kleinen Münzen wird die Größe dagegen eher unterschätzt. Diese Unterschätzung ist beträchtlich und beträgt nach dem Gruppenversuch ungefähr 20% der Münzgröße. Woher kommt diese Unterschätzung, woran liegt das? Bevor wir uns mit dieser Frage beschäftigen, wird im nächsten Abschnitt noch ein weiterer Aspekt der Frage untersucht. „Wie gut“ kann man auch verstehen als „wie genau, wie zuverlässig“. Was ist der Unterschied?

2.4.4. Antwort auf die Versuchsfrage II: Wie genau ist die Schätzung?

Die erste Frage aus dem letzten Abschnitt lautet eigentlich: Wie gut stimmt die subjektive Schätzung mit der objektiven Größe überein? Um diese zu beantworten, wurde der Mittelwert aus den subjektiven Schätzungen pro Münzgröße berechnet und dieser mit den objektiven Größen verglichen. Man könnte die Frage aber auch anders stellen: Wenn man die Größe ein und derselben Münze mehrfach schätzt, kommt dabei immer derselbe Wert heraus oder ist diese Schätzung mit einer Unsicherheit, einer Ungenauigkeit behaftet? Hier kommt der zweite statistische Kennwert ins Spiel, die Streuung bzw. der Variationskoeffizient. Dieser gibt an, wie stark die einzelnen Schätzungen vom Mittelwert abweichen, wobei diese mittlere Abweichung wieder auf den Mittelwert bezogen wird, um damit die unterschiedlichen Grundgrößen der Münzen zu berücksichtigen. Abbildung 11 zeigt das Ergebnis für das einzelne Beispiel.

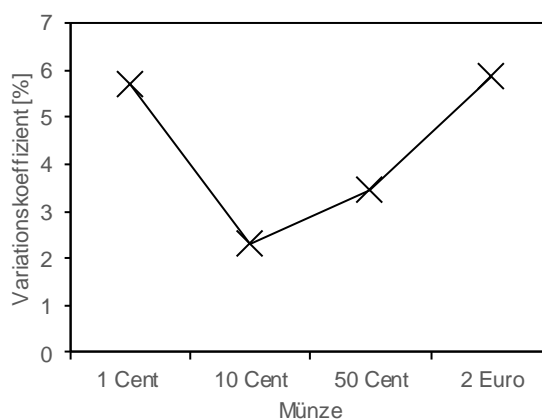


Abbildung 11: Der Variationskoeffizient in Prozent in Abhängigkeit von der Münzgröße für das Beispiel oben.

Man erkennt bei diesem Probanden, dass der Variationskoeffizient bei der 1-Cent-Münze recht groß ist. Dies mag mit der Bildschirmauflösung und dem Mausrad zu tun haben: Beim Einstellen der 1-Cent-Münze führt bereits eine kleine Drehung am Mausrad zu einer relativ großen Veränderung der Pixel für die Darstellung. Man hat also relativ gesehen zur Größe der Münze größere Sprünge bei der Veränderung, was zu einer größeren Streuung bei den Urteilen führt. Rechts erkennt man, dass die Genauigkeit umso geringer wird, je größer die Münze ist. Möglicherweise wird es umso schwieriger, kleine Veränderungen der Größe zu erkennen, je größer eine Münze wird. Dieses Phänomen werden wir später noch genauer kennenlernen.

Wichtig ist an dieser Stelle, dass deutlich geworden ist, dass die Frage nach der Güte der subjektiven Schätzung ganz unterschiedlich beantwortet werden kann und dass verschiedene statistische Kennwerte dazu dienen, diese verschiedenen Aspekte zu beantworten:

- Frage 1: Wie gut ist die Schätzung der Münzgröße im Vergleich zur objektiven Münzgröße? Haben Menschen in diesem Sinne die Münzen in ihrer Vorstellung richtig und genau repräsentiert? Die Antwort wird mit Hilfe des Mittelwerts gegeben im Vergleich zur objektiven Münzgröße.
- Frage 2: Wie genau und präzise ist die Vorstellung und darauf basierende Schätzung der Münzgröße? Wenn man mehrfach schätzt, ist die Schätzung jeweils identisch? Die Antwort auf diese zweite Frage erfolgt mit Hilfe des Variationskoeffizienten oder der Streuung bei mehrfacher Beurteilung einer Münze.

Je nach Fragestellung und Bedeutung der Fragen können auch weitere Kennwerte oder Vergleiche sinnvoll sein. Oft merkt man auch erst bei der Überlegung, wie man eine Frage mit Hilfe von statistischen Berechnungen beantworten kann, welche Aspekte eine solche Frage beinhalten kann. Man berechnet also nicht einfach die typischen Kennwerte, sondern überlegt

vielmehr, durch welche Art von Berechnung man seine Fragestellungen am besten beantworten kann.

Wenn man dann eine Antwort auf seine Frage hat (z.B. „kleine Münzen werden unterschätzt, große realistisch eingeschätzt“), ergibt sich daraus meist die nächste Frage nach der Ursache des Effekts. Woran liegt das? Wir hatten dies am Anfang am Beispiel der Strichlängenschätzung bereits diskutiert. Im nächsten Abschnitt wird dies für das Ergebnis des Größenschätzversuchs noch vertieft.

2.5. Diskussion

2.5.1. Was bedeutet das Ergebnis?

Im Gruppenversuch hatte sich gezeigt, dass die subjektive Einschätzung der Münzgrößen systematisch von der objektiven Größe abweicht. Systematisch heißt dabei, dass die Münzgröße umso stärker unterschätzt wird, je kleiner die Münzen sind. Was wir bisher nicht so deutlich beschrieben hatten, war, dass die subjektive und objektive Größe grundsätzlich aber sehr gut übereinstimmen: Die objektiv kleinste Münze war auch subjektiv am kleinsten geschätzt worden, die 10-Cent-Münze als größer, die 50-Cent-Münze als wiederum größer und die 2-Euro-Münze am größten. Die objektive Reihenfolge der Größe von „klein – größer – noch größer – ganz groß“ fand sich also subjektiv auch wieder. Dies Ergebnis scheint fast trivial zu sein, zeigt aber, dass wir in unserer Vorstellung die Größe durchaus richtig repräsentiert haben, zumindest was die Rangordnung oder Reihenfolge der Münzgrößen betrifft.

Das oben zuerst beschriebene Ergebnis bedeutet vor diesem Hintergrund, dass wir nicht nur die Reihenfolge, sondern auch den Größenunterschied in unserer Vorstellung repräsentiert haben, allerdings mit einem systematischen „Fehler“. Unsere Vorstellung bildet die objektiv messbare, physikalisch Größe also nicht 1:1 ab, sondern verzerrt diese systematisch, wobei der Zusammenhang zwischen objektiver und subjektiver Größe recht gut durch eine Gerade zu beschreiben ist. Es handelt sich also um einen linearen Zusammenhang.

Wenn dieses Ergebnis nicht nur für Münzen, sondern auch für andere Objekte (Stifte, Taschentücher, Bücher, Autos usw.) gelten würde, hätten wir viel gelernt darüber, wie die menschliche Wahrnehmung und die Vorstellung funktionieren. Wir wüssten, dass Menschen nicht nur unterschiedliche Größen in der richtigen Relation zueinander („größer – kleiner“) wahrnehmen, sondern auch die Größenunterschiede repräsentiert haben, allerdings mit einem Verzerrungsfaktor. Wenn wir diesen kennen, können wir vorhersagen, wie groß ein bestimmtes Objekt aus der Vorstellung heraus eingeschätzt wird, da wir aus der objektiven Größe mit Hilfe des Verzerrungsfaktors die subjektive Größe berechnen könnten. Wir wüssten damit, wie groß dieses Objekt in der Vorstellung ist.

- Die Wissenschaft, die sich mit der Abbildung physikalischer Größen in der Wahrnehmung und Vorstellung des Menschen beschäftigt, wird „Psychophysik“ genannt.

Aber sagt dieses Ergebnis tatsächlich, dass der subjektive Fehler bei der Größenschätzung umso größer wird, je kleiner Objekte sind? Die Grundlage für diese Aussage war, dass im Versuch die Größe von Objekten variiert wurde und dann festgestellt wurde, dass die Abweichung je nach Größe systematisch unterschiedlich war. Aber Achtung! Es gibt ein Problem mit dieser Versuchsanordnung. Denken Sie bitte kurz darüber nach, ob Sie selbst darauf kommen, und lesen Sie erst dann weiter!

Das Problem ist, dass sich die vier Münzen nicht nur in der Größe, sondern auch im Wert unterscheiden. Die kleinste Münze ist am wenigsten wert, die größte am meisten. Man könnte deshalb mit gleichem Recht auch behaupten, dass die Größe von Objekten, die wenig wert sind,

stark unterschätzt wird, und die Größe umso realistischer eingeschätzt wird, je mehr die Objekte wert sind. Ist es der Wert oder die Größe der Münze, was für die Unterschätzung verantwortlich ist? Aufgrund unseres Versuchs können wir das nicht unterscheiden. Diese beiden Eigenschaften der untersuchten Objekte sind „konfundiert“, man kann sie nicht voneinander trennen. Eigentlich wollten wir nur eine Eigenschaft, die Größe, untersuchen. Weil wir diese aber an konkreten Objekten, den Münzen, untersucht haben, geht mit der Größe gleichzeitig der Wert einher. Man kann dann nicht entscheiden, welche der beiden Eigenschaften ausschlaggebend war und gewirkt hat.

Der einzige Ausweg aus diesem Dilemma ist ein weiterer Versuch, ein Entscheidungsexperiment. Wie würden Sie dieses planen? Was würden Sie variieren? Bevor wir dies im nächsten Kapitel beschreiben, sollten wir auch noch darüber nachdenken, was wir bei der ersten Untersuchung methodisch gelernt haben. Darum geht es im nächsten Abschnitt.

2.5.2. Was wurde hier eigentlich gemacht? Die Methodik

In dem Versuch wurden systematisch Reize dargeboten, die sich in der Größe (und dem Wert) unterschieden. Die Versuchsfrage war, in welcher Weise die subjektive Größeneinschätzung von der objektiven Größe abhängt. Wie in der Einleitung beschrieben, ging es damit letztlich darum, eine Beziehung der Form „S -> O -> R“ zu beschreiben. Die Reize oder Stimuli waren die vier unterschiedlichen Münzgrößen, auf die Sie als Organismus mit Hilfe einer Größeneinschätzung reagiert haben. Um diese Beziehung zu beschreiben, haben wir systematisch die Reize variiert und die Reaktionen aufgezeichnet. In der Methodik bezeichnet man die Reize, die man im Experiment variiert, als „unabhängige Variable“ (UV) und die Reaktionen, die man misst, als „abhängige Variablen“ (AV). Damit möchte man deutlich machen, dass die UV die Größe ist, die wirken und das Verhalten verändern soll. Die AV, also das Verhalten, soll dagegen davon abhängen, welche Ausprägung der UV vorhanden ist. Diese beiden Begriffe sollten Sie sich unbedingt merken!

- UV oder unabhängige Variable: Die Einflussgröße, die im Experiment systematisch variiert wird, und deren Einfluss auf das Verhalten untersucht werden soll.
- AV oder abhängige Variable: Der Aspekt des Erlebens oder Verhaltens, der sich verändern soll und der im Experiment gemessen wird.

Um die Beziehung „S -> O -> R“ zu beschreiben, wird das Experiment in der Form

- $UV \rightarrow V_p \rightarrow AV$

Durchgeführt (V_p steht hier für „Versuchsperson“). Was man damit zeigen möchte, lässt sich auch beschreiben als

- $AV = f(UV)$

Das bedeutet, wir vermuten, dass die abhängige Variable eine Funktion (f) der unabhängigen Variablen ist, also systematisch von dieser abhängt. In unserem Beispiel hatten wir eine lineare Funktion gefunden, mit der wir den Zusammenhang zwischen objektiver Münzgröße (UV) und subjektiver Einschätzung der Münzgröße (AV) beschreiben konnten. Es kann aber auch jede andere Art von Zusammenhang sein. Darauf werden wir noch genauer eingehen. Auch bei Abbildungen findet sich diese Denkweise wieder: Auf der x-Achse werden in der Regel die unterschiedlichen Abstufungen der UV aufgetragen, im Beispiel also die Münzgröße. Auf der y-Achse werden die Werte der AV, also im Beispiel die geschätzten Münzgrößen dargestellt. Wir sind daran gewöhnt, Grafiken in dieser Weise zu interpretieren: Wie verändern sich die Werte

auf der y-Achse in Abhängigkeit von den Stufen der x-Achse? Dies entspricht genau dem Zusammenhang, der im Experiment untersucht werden soll.

Das Experiment erlaubt diesen Schluss aber nur, wenn tatsächlich nur die UV in ihren Ausprägungen variiert wurde und alle anderen Bedingungen konstant gehalten wurden. Dieses Prinzip hatten wir in unserem Experiment verletzt, weil neben der Größe auch der Wert der Münzen variiert wurde. Das führt dazu, dass wir nicht wissen, ob Größe oder Wert für die systematische Veränderung der AV verantwortlich sind. Wir müssen daher versuchen, die wesentliche Wirkgröße zu isolieren (s. nächstes Kapitel). Es könnte aber noch eine Vielzahl anderer Störgrößen vorliegen. Sie haben vielleicht selbst gemerkt, dass Sie im Laufe des Versuchs müde und unkonzentrierter geworden sind. Hätten wir am Anfang des Versuchs nur die großen Münzen, am Ende nur die kleinen Münzen einschätzen lassen, hätte der Unterschied in der Einschätzung auch an der nachlassenden Konzentration oder zunehmenden Müdigkeit liegen können. Um derartige Zeiteffekte zu vermeiden, hatten wir Sie einerseits gebeten, eine Pause einzulegen, wenn Sie müde oder unkonzentriert werden. Andererseits hatten wir in jedem Durchgang alle Münzen, aber jeweils in zufälliger Reihenfolge gegeben. Über den Versuch hinweg, müsste damit die Müdigkeit oder nachlassender Konzentration bei allen Münzgrößen gleichmäßig wirken. Man nennt dieses Prinzip „Kontrolle von Störvariablen“.

- Kontrolle von Störvariablen: Im Experiment wird versucht, alle Einflussgrößen, die nicht systematisch als UV variiert werden, auszuschließen oder gleichmäßig auf alle Stufen der UV zu verteilen.

Durch die Kontrolle von Störvariablen versucht man also genau, solche Konfundierungen wie die zwischen Größe und Wert in unserem Experiment zu vermeiden und sicherzustellen, dass nur die unterschiedlichen Stufen der UV gewirkt haben können. Je besser dies gelingt, umso eindeutiger ist ein gefundener Effekt zu interpretieren. Wir werden auf entsprechende Kontrolltechniken im Lauf dieser Übung noch genauer eingehen. Doch jetzt wollen wir versuchen, eine Entscheidung zu treffen: Ist es die Münzgröße oder der Wert, was zu der Unterschätzung führt?

2.6. Versuch Größeneindruck II: Wert oder Größe?

2.6.1. Ziel des Experiments

Wir wollen jetzt mit Hilfe eines Experiments entscheiden, ob der Wert oder die Größe für die Unterschätzung verantwortlich sind. Wir müssen also diese beiden Einflussgrößen voneinander trennen bzw. sie isoliert voneinander untersuchen. Die einfachste Möglichkeit ist es, den Wert zu eliminieren und nur die Größe zu variieren. Um dies zu tun, wird Ihnen beim nächsten Versuch (Größeneindruck II) jeweils ein Kreis mit der Größe der entsprechenden Münze kurz gezeigt, dann gelöscht und dann sollen Sie einen weiteren Kreis so groß einstellen wie der Kreis, den Sie vorher kurz gesehen haben. Denken Sie dabei nicht an die Münzen! Versuchen Sie, sich ganz auf den Kreis zu konzentrieren, den Sie zuerst gesehen haben und diesen möglichst gut wieder einzustellen.

Wenn es wirklich der Wert der Münzen ist, die den Effekt hervorgerufen hat, sollten dieser Effekt jetzt nicht mehr auftreten. Sehen wir also mal, was bei Ihnen passiert.

2.6.2. Instruktion

Den Versuch finden Sie im Programm bei dem Menüpunkt „Größeneindruck II“. Bevor Sie den Versuch starten, lesen Sie bitte die folgende Instruktion aufmerksam durch! Sie wissen, es ist

wichtig, dass Sie eine ganz klare Vorstellung davon haben, um was es im Versuch geht und wie Sie an den Versuch herangehen sollen.

In dem Versuch werden Ihnen vier unterschiedlich große Kreise jeweils kurz gezeigt. Merken Sie sich jedes Mal die Größe des Kreises. Nach einer kurzen Pause haben Sie dann die Möglichkeit, einen weiteren Kreis mit Hilfe des Mauseisens in der Größe zu verändern. Stellen Sie sich jetzt möglichst gut den Kreis vor, den Sie sich gemerkt haben und stellen Sie die Kreisgröße so ein, dass er diesem gemerkten möglichst gut entspricht. Pendeln Sie dazu um den besten Wert hin und her bis Sie den Punkt gefunden haben, der am besten zu ihrer Vorstellung passt. Dann drücken Sie die linke Maustaste. Am Ende wird eine Tabelle gezeigt, mit der Sie die Werte in die Excel-Tabelle übertragen können. Jeder der vier Kreise wird wieder viermal gezeigt.

Bitte sorgen Sie dafür, dass Sie beim Versuch ungestört sind und konzentriert arbeiten können. Wir zeigen den Kreis nur achtmal, sodass der Versuch nicht lange dauert.

Ganz wichtig ist es für Sie, dass Sie jetzt nicht an die Hypothesen und Ergebnisse des ersten Versuchs denken. Versuchen Sie also wirklich, sich nur auf ihr Gedächtnis zu verlassen und denken Sie nicht darüber nach, dass Sie etwas ausgleichen oder korrigieren müssten. Es soll ja um ihre unmittelbare Vorstellung gehen!

Und hier geht es zum Versuch:

- <https://cgi-psych.rz.tu-bs.de>

2.6.3. Auswertung und Ergebnis

Übertragen Sie die Tabelle aus dem Programm nach Excel, dort in den Reiter „Größeneindruck 2“. Markieren Sie auch hier in Excel die Tabelle und fügen Sie dann die Daten aus dem Programm ein mit „Inhalte einfügen – Text“. Die Berechnungen sind dieselben wie in der ersten Tabelle. Abbildung 12 zeigt das Ergebnis für den Probanden, dessen Beispiel auch oben verwendet wurde. Bei ihm fällt das Ergebnis sehr eindeutig aus:

- Alle Abweichungen liegen im negativen Bereich und die Abweichungen sind deutlich kleiner als bei den Münzen.

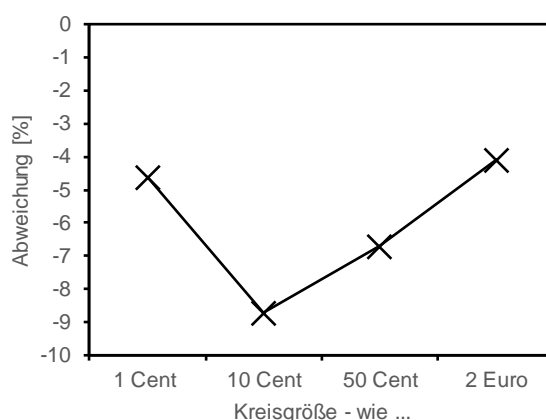


Abbildung 12: Ergebnis des zweiten Teils des Selbstversuchs. Dargestellt sind die Abweichungen von der Größe der Kreise (die der Größe der entsprechenden Münzen entsprach) in Prozent.

Das Entscheidungsexperiment hat also ein klares Ergebnis gebracht: Der Wert der Münze scheint dafür verantwortlich zu sein, dass die Größe unterschätzt wird. Die eigentliche unabhängige Variable, die für den Effekt verantwortlich ist, ist also der unterschiedliche Wert der

Münzen, nicht aber deren Größe. Die psychologische Gesetzmäßigkeit, die wir damit gefunden haben, heißt also:

- „Objekte mit geringem Wert werden in der Größe eher unterschätzt.“

2.6.4. Diskussion und Zusammenfassung

- „Objekte mit geringem Wert werden in der Größe eher unterschätzt.“

Wenn man das so formuliert, fällt sofort auf, dass dieser Satz eigentlich viel zu allgemein ist für unser Experiment. Es könnte ja sein, dass das nur für den Euro gilt, nicht aber für andere Währungen. Oder nur in Deutschland, weil die Deutschen viel Wert auf Geld legen. Oder nur für Studenten, weil da jeder Euro zählt. Oder nur für Psychologiestudenten. Um zu zeigen, dass das Ergebnis so allgemein, wie es oben formuliert wurde, gilt, müsste man das Experiment also wiederholen, mit anderen Studenten, anderen Deutschen, anderen Menschen. Angesprochen ist hier ein generelles Problem von empirischen Wissenschaften: Die Experimente geschehen immer unter bestimmten Bedingungen mit bestimmten Probanden. Eigentlich kann man damit allgemeine Sätze, die für alle Menschen gelten sollen, nicht beweisen. Um einem Beweis näher zu kommen, sollte man also Experimente mit anderen Personen wiederholen, also replizieren, um zu sehen, ob es auch für diese gilt. Und wenn man das Ergebnis so allgemein formuliert hat, könnte man versuchen, die Gültigkeit auch damit zu prüfen, dass man z.B. andere Objekte untersucht. Führen Sie den Versuch doch mal mit Diamanten durch!

Bei dem Entscheidungsexperiment gab es natürlich auch noch andere Probleme. Sie durften die Münzen direkt vor dem Versuch nicht mehr anschauen, sondern sollten deren Größe aus dem Langzeitgedächtnis, aus ihrer Vorstellung heraus schätzen. Die Kreise bei Größeneindruck II hatten Sie direkt vorher gesehen und sollten sie nach einer sehr kurzen Pause einschätzen. Es war also möglicherweise noch in ihrem Kurzzeitgedächtnis vorhanden. Vielleicht liegt die bessere Schätzung ja daran, dass ein anderes Gedächtnissystem beteiligt ist. Man müsste den Münzversuch also nochmal so wiederholen, dass unmittelbar vor der Schätzung die jeweilige Münze gezeigt wird, man sich deren Größe noch einmal einprägen kann, und man sie dann nach einer kurzen Pause einschätzt. Auch das können Sie einfach selbst ausprobieren, indem Sie den Versuch entsprechend wiederholen.

Der fehlende Effekt bei den Kreisen könnte aber auch noch andere Ursachen haben als den Wert der Münze. Eine Münze ist aus Metall, hat eine bestimmte Dicke, eine bestimmte Farbe, eine Beschriftung. Vielleicht würde die Unterschätzung auch auftreten, wenn man entsprechende Metallscheiben verwendet. Oder vielleicht lag es an der Farbe der 1-Cent-Münze? Diese Überlegungen zeigen, wie sorgfältig man nachdenken muss, wenn man die Abstufungen einer UV auswählt. Untersucht man damit tatsächlich das, was man untersuchen will? Oder sind doch andere Eigenschaften damit verbunden oder konfundiert?

Sie haben in diesem ersten Teil eine ganze Menge methodisch gelernt. Psychologen beschäftigen sich damit, wie Reize das Verhalten verändern und möchten wissen, was dabei im Organismus, im Menschen, vor sich geht. Im Experiment werden deshalb gezielt verschiedene Reize, die als Stufen der unabhängigen Variablen bezeichnet werden, variiert. Man misst als abhängige Variable, wie sich die Probanden verhalten, und schließt dann auf die zugrundeliegenden psychischen Prozesse. Durch geschickte Variation der unabhängigen Variablen kann man den wirklich relevanten Eigenschaften der unabhängigen Variablen auf die Spur kommen.

Auch inhaltlich haben Sie etwas Wichtiges gelernt: Für die Wahrnehmung und die Vorstellung sind nicht nur die physikalischen Eigenschaften von Reizen wichtig, sondern deren Wert und Bedeutung für Sie als Menschen. Wert und Bedeutung sind nicht nur eine eigene Eigenschaft

von Objekten, sondern verändern sogar unsere Vorstellung von physikalischen Eigenschaften. Wertvolle Objekte werden größer eingeschätzt, Objekte mit geringem Wert zu klein. Unsere Welt wird also durch Werte geprägt.

3. Versuch Streckenmitteln: Das psychische Maß – ist subjektives Erleben messbar?

Lesen Sie einmal die Überschrift bewusst durch. Verstehen Sie, was diese bedeutet? Ich habe es beim ersten Lesen nicht verstanden. Was ist ein „psychisches Maß“? Was bedeutet es, dass subjektives Erleben „messbar“ sein soll? Natürlich kann ich fragen, wie es jemandem geht. „Gut, danke der Nachfrage“, wäre eine mögliche Antwort. Also, das subjektive Erleben ist „gut“. Habe ich damit schon gemessen? Was ist eine Messung?

Um diese Fragen geht es in dem folgenden Kapitel. Dazu gibt es im ersten Abschnitt zunächst etwas Theorie, nämlich die Frage, was eine Messung überhaupt ist.

3.1. Was ist eine Messung oder Metrik?

Im letzten Kapitel haben wir gesehen, dass die Wahrnehmung der Größe unter bestimmten Bedingungen auch vom Wert des betreffenden Gegenstands abhängig ist. Das 1-Cent-Stück mit einer gemessenen Größe von 16.3 mm wurde von den Probanden mit 12.2 mm (s. Abbildung 12) deutlich kleiner eingeschätzt. Bei dem Entscheidungsexperiment, bei dem wir einen Kreis mit der Größe von 16.3 mm eingeschätzt hatten, war die geschätzte Größe dagegen ziemlich identisch mit der objektiv gemessenen (16.5 mm). Wenn man das gegenüberstellt, heißt das, dass etwas gleich Großes (1-Cent-Stück und gleich großer Kreis) subjektiv unterschiedlich groß ist (12.2 und 16.5 mm). Das sieht zunächst nach einem Fehler aus. Unsere Überlegungen haben aber gezeigt, dass vermutlich der Wert die Größenschätzung beeinflusst. Es handelt sich hier also um keinen „Fehler“, sondern um eine Eigenschaft unseres Erlebens, um ein psychisches Phänomen. Wir haben damit den subjektiven Wert einer Münze „gemessen“ – die Unterschätzung wurde immer größer, je geringer der Wert war. Wir haben also beschrieben, wie stark der Wert die Größenschätzung beeinflusst.

Was heißt messen? Formal geht es darum, dass wir bestimmten Eigenschaften Zahlen zuordnen. Ein Kreis einer bestimmten Größe wird gemessen, indem wir ein Lineal anlegen und die entsprechende Länge dort ablesen – als Zahlen. Die Beziehungen zwischen den Zahlen (2 ist größer als 1) geben die Beziehungen zwischen den Objekten wieder (die 2-Euro-Münze ist größer als die 50-Cent-Münze). Wenn wir die Unterschätzung der Münzgröße als Maß für den subjektiven Wert ansehen, haben wir diesen gemessen, dem Wert also eine Zahl zugeordnet. Die 10-Cent-Münze hat einen geringeren Wert als die 2-Euro-Münze, weil sie stärker unterschätzt wird, während die 2-Euro-Münze annähernd realistisch eingeschätzt wird. Wir haben damit eine bestimmte psychische Eigenschaft, den Wert, auf eine bestimmte Art und Weise gemessen.

In der Physik sind die wesentlichen Eigenschaften oder Dimensionen, die gemessen werden, die Größe (oder Länge), das Gewicht und die Zeit. Diese werden in bestimmten Einheiten, z.B. Meter, Kilogramm und Sekunden, gemessen. Jeder Gegenstand ist physikalisch mit diesen drei Eigenschaften zu bestimmen: Wie groß ist er? Wie schwer ist er? Wann war der Gegenstand an diesem Ort? In diesen drei Dimensionen dient die Messung dazu, Relationen zwischen den Gegenständen anzugeben: „Die 2-Euro-Münze ist größer und schwerer als die 10-Cent-Münze.“ „Zuerst wurde die 2-Euro-Münze gezeigt, dann die 10-Cent-Münze.“ Die einfachste Relation ist „gleich“ vs. „ungleich“, etwas komplexer ist „kleiner“ – „gleich“- „größer“. Auf weitere Relationen werden wir später noch eingehen.

Wichtig ist außerdem, dass diese Messung unterschiedlich genau geschehen kann. Wir haben die Münzgrößen in Millimeter gemessen, man könnte sie aber auch in Zentimetern angeben. Meterangaben machen bei Münzen dagegen keinen richtigen Sinn. Um Entfernung zwischen zu Hause und der Arbeitsstelle anzugeben, sind dagegen millimetergenaue Messungen wenig hilfreich. Je nach Gegenstand kann demnach unterschiedlich genau gemessen werden.

Damit haben wir definiert, was Messen ist: Die Zuordnung von Zahlen zu bestimmten Eigenschaften von Objekten. Als „Metrik“ bezeichnet man ein Messsystem, das folgende beide Aspekte umfasst:

- Die Eigenschaft oder Dimension, die gemessen wird: die Qualität
- Die Messgenauigkeit oder die noch zu unterscheidenden Differenzen: die Quantität

Die Metrik der Physik hat drei wesentlichen Qualitäten (Größe, Gewicht, Zeit), die in unterschiedlichen Quantitäten gemessen werden (z.B. Meter, Kilogramm, Sekunden). Die Psychologie orientiert sich hier an der Physik und Mathematik. Als Psychologen interessiert uns, welche Qualitäten in der Psychologie existieren: Was sind die unterschiedlichen Erlebnisse, die wir spüren? Damit einher geht die Frage nach der Quantität: Wie genau sind unsere Erlebnisse? Wie ist unsere Empfindungsfähigkeit ausgeprägt?

Das ist mit dem Begriff in der Überschrift gemeint. Wenn man nach dem „psychischen Maß“ fragt, geht es um die Metrik des Psychischen, also: Welche Dimensionen des Erlebens mit welcher Unterscheidungsfähigkeit gibt es? Wie genau können wir was empfinden? Der nächste Abschnitt beschreibt, wie man dies methodisch untersucht.

3.2. Die Bedeutung des „Gleich“-Urteils

In der Physik ist Messen einfach: Man legt einen Maßstab, ein Lineal an. Man misst die 10-Cent -Münze: 16.3 mm. Man misst den Kreis aus dem Entscheidungsexperiment: 16.3 mm. Die gleiche Zahl, also sind die beiden Objekte gleich groß. Für die Physik ja, für uns als Menschen aber nicht. Physikalische Metrik ist also nicht gleich der psychischen Metrik – unser Erleben ist anders als die Physik. Aber hat der Mensch überhaupt eine Metrik? Gibt es eine psychische Metrik?

Wir haben gerade gesehen, dass zu einer Metrik zwei Dinge gehören:

- 1) Die Dimension (Qualität)
- 2) Der Genauigkeitsbereich (Quantität)

Für die Bestimmung beider Aspekte dieser psychischen Metrik ist die Frage „Was ist gleich?“ ganz entscheidend. Die Münze ist nicht gleich dem Kreis, obwohl sie von der Größe her gleich groß sind. Auch wenn man uns als Menschen „Größe“ beurteilen lässt, scheinen wir das nur schwer vom Wert trennen zu können. Der Wert beeinflusst die Größe. Die psychische Dimension der Größe beinhaltet offensichtlich auch irgendwie den Wert. Damit trägt die Frage „Was ist gleich?“ dazu bei, die Dimension genauer zu bestimmen. Die Frage nach der Dimension hatten wir experimentell versucht zu lösen, indem wir ein Entscheidungsexperiment durchführten. Als ein Kreis von der Größe einer 10-Cent-Münze beurteilt werden sollte, war keine Unterschätzung der Größe zu finden. Genau daraus schlossen wir ja, dass der Wert eine Rolle spielt.

Die Frage „Was ist gleich?“ ist aber auch für den Genauigkeitsbereich, die Quantität, ganz entscheidend. Das sieht man auch in der Physik. Wenn Ihr Maßstab nur eine Einteilung nach Zentimetern hat, können Sie eben nur auf Zentimeter genau messen. Je nachdem, welche Einheit Ihre Landkarte hat, können Sie die Weglänge mehr oder weniger genau messen. In der Psychologie ist der Maßstab unser Erleben – und natürlich können wir auch fragen: „Sind zwei Objekte gleich groß?“ oder „Sind zwei Objekte unterscheidbar, unterschiedlich groß?“ Wie groß muss ein Unterschied zwischen zwei Empfindungen sein, damit wir ihn überhaupt merken?

Damit können wir die Frage, ob es auch eine psychische Metrik gibt, beantworten, indem wir die Frage genauer untersuchen: Wann sagen wir „gleich“? In die Untersuchung dieser Frage

steigen wir mit dem nächsten Versuch ein. Im folgenden Abschnitt finden Sie die Anleitung und Instruktion zu diesem Versuch.

3.3. Anleitung zum Versuch „Streckenmitteln“

Unsere Ausgangsfrage des Versuchs heißt: „Wann sagen wir gleich?“ Wir werden das an Strecken überprüfen. Wenn wir eine Strecke mitteln, dann entstehen zwei gleiche Stücke. Wenn Sie das Programm starten, wird Ihnen deshalb eine waagerechte Strecke gezeigt. Diese Strecke soll mit Hilfe einer senkrechten Linie genau gemittelt werden. Dazu können Sie die senkrechte Linie nach links und rechts bewegen, indem Sie die Maus nach rechts und links bewegen.

Um die Mitte zu finden, sollen Sie das „Herstellungsverfahren“ anwenden. Sie kennen das eigentlich schon vom ersten Versuch, dem Größenschätzen. Sie kennen das wahrscheinlich auch aus dem Alltag, zum Beispiel, wenn Sie einen Beamer scharf einstellen. Sie drehen dabei die Linse zunächst grob in den Bereich, indem das Bild scharf wird. Dazu stellen Sie in großen Drehungen schärfer und unschärfer. Wenn Sie den groben Bereich identifiziert haben, machen Sie kleinere Bewegungen und pendeln sich damit auf den schärfsten Bereich ein. Dieses Vorgehen übertragen wir jetzt systematisch auf das Streckenmitteln:

- Gehen Sie mit dem senkrechten Strich ganz nach links, dann von da aus grob in die Mitte. Wiederholen Sie das von rechts.
- Jetzt haben Sie ein grobes Gefühl, wo die Mitte ist.
- Gehen Sie jetzt ein bisschen nach links, bis die linke Hälfte sicher kürzer ist. Dann zurück nach rechts, bis die rechte Hälfte wieder kürzer ist.
- Jetzt haben Sie den Bereich eingegrenzt.
- Jetzt wieder nach links und rechts, aber nicht ganz so weit. Das ist die Feineinstellung.
- Und zum Schluss nur ein ganz kleines bisschen, dann sind Sie genau in die Mitte.

Lassen Sie wieder jedes Lineal weg. Verlassen Sie sich ganz auf Ihren Eindruck. Wir wollen ja nicht physikalisch messen, sondern das psychische Maß untersuchen. Wichtig ist also, was Sie sehen!

Wenn Sie die Mitte gefunden haben, drücken Sie die linke Maustaste. Sie können sich kurz neu konzentrieren und dann mit der nächsten Strecke weitermachen. Ganz am Ende erscheint eine Tabelle, in der alle ihre Mittelungen enthalten sind. Diese sollen Sie wieder nach Excel kopieren.

Noch ganz kurz als Hintergrund: Es wird wieder verschieden lange Strecken geben. Da es eine Rolle spielen kann, in welcher Reihenfolge diese auftauchen, werden Sie eine zufällige Reihenfolge von langen und kurzen Strecken erhalten. Wir werden auch jede Strecke insgesamt sieben Mal präsentieren. Ein Aspekt von Genauigkeit ist ja, dass bei mehrfacher Messung dasselbe Ergebnis entsteht. Das werden wir untersuchen, indem wir die Messung wiederholen.

Auch hier ist es natürlich wieder wichtig, dass Sie möglichst konzentriert und ungestört arbeiten und sich nicht ablenken lassen. Wenn Sie merken, dass die Konzentration nachlässt, machen Sie bitte eine Pause. Alles klar? Dann rufen Sie jetzt bitte das Programm „Streckenmitteln“ auf und beginnen Sie.

Und hier geht es zum Versuch:

- <https://cgi-psych.rz.tu-bs.de>

3.4. Auswertung des Versuchs

3.4.1. Subjektive Mitte und Unsicherheitsintervall

Das Programm hat die verschiedenen Strecken schon in die richtige Reihenfolge gebracht. Es hat die Abweichungen von der „wahren“ Mitte bestimmt. Negative Werte bedeuten, dass Sie einige Pixel nach links von dieser physikalischen Mitte abweichen. Positive Werte sind eine Abweichung nach rechts. Die Werte, die Sie sehen, sind Pixel, also Punkte auf dem Bildschirm. Kopieren Sie diese Werte in Ihre Excel-Tabelle (s. oberer Teil von Tabelle 6).

Unter dieser Tabelle finden Sie eine zweite, in der die Streckenlänge der halbierten Strecken stehen (Einheit: Pixel). Bilden Sie für jede Spalte (und damit jede Strecke) jetzt den Mittelwert. Machen Sie das in der Zeile unterhalb der Streckenlängen (unterer Teil in Tabelle 6). Sie können dazu wieder die „=Mittelwert(E7:E11)“-Funktion verwenden.

Tabelle 6: Beispiel für die Ergebnisse des Versuchs und erste Auswertung.

Streckenlänge	293	263	230	200	170	143	110
1. Durchgang	9	13	6	7	5	6	2
2. Durchgang	6	6	8	-1	2	4	6
3. Durchgang	-8	-4	7	2	2	1	1
4. Durchgang	3	3	9	6	2	2	2
5. Durchgang	4	3	6	2	5	-2	4
Streckenlänge	293	263	230	200	170	143	110
Mittelwert	3	4	7	3	3	2	3

Der Mittelwert aus den fünf Durchgängen ist eine gute Schätzung dafür, inwieweit Ihre „subjektive Mitte“ von der physikalischen Mitte abweicht. Ist der Mittelwert negativ, dann ist Ihre subjektive Mitte etwas nach links verschoben. Ist er positiv, liegt sie etwas mehr rechts.

- Subjektive Mitte SM: Abweichung von der physikalischen Mitte

Bei dem Probanden im Beispiel liegen die Mittelwerte alle im leicht positiven Bereich, also etwas nach rechts verschoben. Die größte Abweichung der 7 Mittelwerte sind 7 Pixel bei einer mittleren Strecke, was immer noch sehr nahe an der „wahren“ Länge ist, also der physikalisch gemessenen Mitte. Die Frage des Versuchs war, wie genau unsere Wahrnehmung eigentlich ist, denn die Genauigkeit war das eine Bestimmungsstück einer psychischen Metrik. Wenn Sie sich jetzt zunächst die Werte pro Strecke anschauen, werden Sie feststellen, dass es „die“ Mitte nicht gibt. Die Werte in einer Spalte sind recht unterschiedlich. Umgekehrt heißt das, dass „Gleich“ für einen ganzen Bereich von Werten gilt. Das haben Sie sicherlich auch beim Streckenmitteln gemerkt: Man nähert sich von der einen Seite, dann von der anderen, stellt dann feiner ein, und entscheidet. Aber wie man sieht, streuen die Urteile von Durchgang zu Durchgang. Inhaltlich heißt das eben, dass man in diesem Bereich nicht gut unterscheiden kann. Die Werte der sieben Durchgänge sind alle mehr oder weniger gute Beispiele für „gleich“. Es gibt also ein Intervall von Werten, in dem man sich unsicher ist:

- Unsicherheitsintervall (UI): Der Bereich von Werten, bei dem Sie jeweils „gleich“ geurteilt haben.

Dieses Unsicherheitsintervall hilft uns damit, die Frage nach der Genauigkeit des Erlebens zu beantworten: Je größer dieser Bereich ist, umso ungenauer ist unser Erleben. Je ähnlicher sich die Werte dagegen sind, umso sicherer konnten wir urteilen und umso genauer ist unser Erleben.

Fassen wir zusammen: Das Gleich-Urteil gilt für unterschiedliche Messwerte. Die Messwerte streuen um die subjektive Mitte SM. Sie bilden also ein Unsicherheitsintervall, womit die Genauigkeit des subjektiven Erlebens beschrieben werden kann.

3.4.2. Varianz und Standardabweichung

Um zu beschreiben, wie groß das Unsicherheitsintervall ist, könnten wir die mittlere Abweichung oder den Variationskoeffizienten berechnen, den wir in Kapitel 2.3 kennengelernt haben. Können Sie sich noch erinnern?

Für die mittlere Abweichung wurde der Absolutwert der Abweichungen der Messwerte vom Mittelwert addiert und durch die Anzahl der Messwerte geteilt.

$$a = \frac{D}{n}$$

Beim Variationskoeffizienten hatten wir die mittlere Abweichung durch den Mittelwert geteilt und mit 100 multipliziert und so eine prozentuale Abweichung erhalten.

$$a\% = \frac{a}{M} * 100$$

In der psychologischen Literatur werden meist noch zwei andere Maße verwendet, Varianz und Standardabweichung. Bei der Varianz berechnet man nicht den Absolutwert der Abweichungen vom Mittelwert, sondern quadriert die Abweichungen vom Mittelwert. Man teilt dies nicht durch die Anzahl der Messwerte, sondern durch die Anzahl der Messwerte minus 1. Das hat etwas mit höherer Statistik zu tun, weil man mit der Varianz die Werte nicht nur in der Stichprobe, sondern in einer Population schätzen will. Aber das ist hier nicht so entscheidend. Hier die Formel:

$$sd^2 = \frac{D^2}{n - 1}$$

Dabei ist D^2 die Summe der Abweichungsquadrate, n die Anzahl der Messwerte. Warum die Varianz als „ sd^2 “ bezeichnet wird, erklären wir gleich. Um die Berechnung nachzuvollziehen, erstellen Sie eine Tabelle, in der die Abweichungsquadrate enthalten sind. Tabelle 7 zeigt dies für das Beispiel. Sie können für die Berechnung wieder die Formelfunktion von Excel verwenden. Um einen Wert zu berechnen, geben Sie z.B. ein: „=Potenz(E7-E\$18;2)“, wenn ihr Abweichungswert in der Zelle E7 steht und der Mittelwert in Zelle 18. Diese Zelle mit der Formel können Sie dann einfach in die anderen Zellen der Tabelle kopieren. Excel passt dann Zeilen und Spalten an. Bei allen 5 Abweichungswerten in einer Spalte soll ja immer die Differenz zu dem Mittelwert dieser Spalte gerechnet werden. Deshalb wurde bei „E\$18“ das „\$“ vor die 18 gesetzt. Wenn Sie diese Formel kopieren, passt Excel die Zeilen und Spalten an bis auf die Zeile, vor der dieses „\$“ steht.

Tabelle 7: Abweichungsquadrate für das Beispiel und Berechnung der Varianz

Quadierte Abweichung	293	263	230	200	170	143	110
1. Durchgang	38	77	1	14	3	14	1
2. Durchgang	10	3	1	18	1	3	9
3. Durchgang	117	67	0	1	1	1	4
4. Durchgang	0	1	3	8	1	0	1
5. Durchgang	1	1	1	1	3	18	1
Varianz	42	38	2	11	3	9	4
Standardabweichung	6	6	1	3	2	3	2

Wenn Sie dies für Ihre Messwerte getan haben, können Sie die Varianz berechnen, indem Sie die Werte aufsummieren (Funktion „=Summe()“) und dann durch 4 teilen ($n-1$). Die entsprechenden Werte sehen Sie in Tabelle 7 in der vorletzten Zeile. Die Varianz hat aber einen Nachteil: Da wir die Abweichungswerte quadriert haben, ist die Einheit der Varianz nicht Pixel, sondern Pixel zum Quadrat. Um dies auf dieselbe Einheit wie die Messwerte zu bringen, wird

einfach die Wurzel aus der Varianz gezogen. Dieser Kennwert heißt „Standardabweichung“ und wird gerne mit „sd“ abgekürzt (deshalb auch die Abkürzung „sd²“ für die Varianz).

$$sd = \sqrt{sd^2}$$

Dies finden Sie in der letzten Zeile der Tabelle. Um die Wurzel zu berechnen, benutzen Sie die Formel „=Wurzel(E26)“, wenn die Varianz in Zelle E26 steht. In der Tabelle wurden die Nachkommastellen gerundet. In Excel geht das einfach, indem Sie die Zellen formatieren (Zelle markieren und rechte Maustaste drücken, dann „Zelle formatieren“ auswählen). Bei dem Reiter „Zahlen“ findet man die Kategorie „Zahl“ und dort können Sie rechts die Anzahl der Dezimalstellen einstellen. Wählen Sie hier „0“, um auf ganze Zahlen zu runden.

Sie haben jetzt zwei neue, sehr häufig verwendete Kennwerte kennengelernt, mit denen man die Streuung von Messwerten beschreiben kann. Wozu brauchen wir diese? Wir hatten ja bei den Ergebnissen gesehen, dass wir die Strecken nicht immer gleich geteilt hatten, sondern dass sich die Messwerte von Durchgang zu Durchgang unterscheiden. Inhaltlich bedeutet das, dass unser Gleich-Urteil nicht ein bestimmter Punkt ist, sondern dass es sich um einen bestimmten Bereich handelt. Physikalisch ungleiche Strecken erscheinen uns im Erleben gleich. Oder: Unser Erleben ist nicht so genau, wie wir es mit Methoden der Physik messen können. Die Streuung soll als Maß dafür dienen, wie (un-) genau unser Erleben ist. Dies wird im folgenden Abschnitt beschrieben.

3.4.3. Berechnung von Unsicherheitsintervall und Unterschiedsschwelle

Wir wollen beschreiben, wie groß der Bereich ist, der für unser Erleben gleich erscheint. Der Mittelwert gibt die Mitte dieses Bereichs an. Von dort aus kann man hin zum „Größer“ und in die andere Richtung zum „Kleiner“ gehen. Als Unsicherheitsintervall (UI) wird deshalb der Bereich definiert, der vom Mittelwert minus der Standardabweichung bis zum Mittelwert plus der Standardabweichung reicht, um diese beiden Richtungen einzubeziehen. Als Formel:

$$UI = m \pm sd$$

In dem Beispiel oben war der Mittelwert für die längste Strecke zum Beispiel -11 Pixel. Die Standardabweichung war 12 Pixel. Das Unsicherheitsintervall reicht also von -23 Pixel bis 1 Pixel. Solange die Teilung der Strecke in diesem 24 Pixel breiten Bereich liegt, hat der Proband im Beispiel den Eindruck, dass die beiden Teilstrecken gleich lang sind.

Auch dies sollten Sie in Excel berechnen. Die untere Grenze des UI bestimmen Sie, indem Sie vom Mittelwert die sd abziehen. Für die obere Grenze addieren Sie die sd.

Wie weit muss man sich vom Mittelwert entfernen, damit der Eindruck „länger“ oder „kürzer“ entsteht? Nach der Formel lautet die Antwort: „Jeweils um eine Standardabweichung.“ Dies bezeichnet man auch als Unterschiedsschwelle US. Damit ist die Berechnung der US in der Tabelle in Excel einfach: Die US ist einfach die Standardabweichung! Abbildung 13 zeigt die Unterschiedsschwelle für die verschiedenen Streckenlängen für das Beispiel.

Man erkennt, dass die Schwelle umso größer ist, je länger die Linie ist. Dieser Zusammenhang wird auch durch die gestrichelte Linie verdeutlicht. Diese wurde mit Hilfe eines statistischen Verfahrens, der linearen Regression, berechnet. Wie das genau funktioniert, werden sie sicherlich in Ihren Statistikkursen noch kennenlernen. Wenn man in Excel die Datenpunkte markiert und dann die rechte Maustaste drückt, erscheint in dem Menü ein Punkt „Trendlinie hinzufügen“. Wenn Sie diesen auswählen, wird eine solche Gerade gezeichnet.

Doch zurück zum Ergebnis: Die Unterschiedsschwelle ist für kurze Linien kleiner als für lange Linien. Etwas allgemeiner ausgedrückt, bemerkt man bei kleinen Gegenständen bereits kleine

Veränderungen, während sich große Gegenstände stärker verändern müssen, damit man diese Veränderung bemerkt. Wenn Sie eine neue Pflanze im Blumentopf eingesät haben und das erste Grün erscheint, können Sie am Anfang fast beim Wachsen zuschauen. Ist die Pflanze aber erst mal groß, muss sie schon stark wachsen, damit Sie es bemerken.

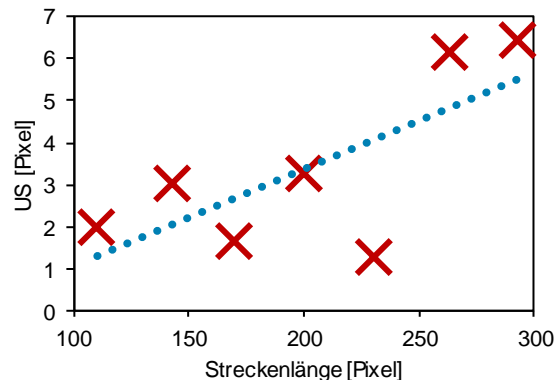


Abbildung 13: Unterschiedsschwelle in Abhängigkeit von der Länge der Linie in Pixeln. Dargestellt sind die Werte aus dem Beispiel. Die gestrichelte Linie zeigt die Regressionsgerade.

Wir haben hier eine ganz wichtige psychologische Gesetzmäßigkeit entdeckt. Es besteht ein Zusammenhang zwischen der Reizgröße und der Größe der Veränderung des Reizes, damit die Veränderung bemerkt wird. Die Unterschiedsschwelle ist also nicht absolut, sondern hängt ab von der Größe des Reizes.

In unserem Versuch bedeutet das, dass bei der kürzesten Strecke schon ein Pendeln von 3 Pixeln zu einem „länger“ oder „kürzer“ Eindruck führt. Bei der längsten Strecke erscheint bei einem solchen Pendeln noch alles gleich und nicht unterscheidbar. Hier benötigt man schon 12 Pixel oder mehr, damit man einen Unterschied bemerkt.

Diese Abhängigkeit der Unterschiedsschwelle von der Reizgröße begegnet Ihnen täglich. Tagsüber können Sie auch bei klarem Himmel die Sterne nicht sehen, weil deren Lichtstärke zu klein ist gegenüber der Helligkeit, die durch die Sonne entsteht. Nachts, bei geringer Ausprägung des Hintergrundlichts, leuchten die Sterne dagegen wunderschön. Auch für die anderen Sinne gilt diese Gesetzmäßigkeit: In der Bibliothek, einem ruhigen Raum, stört jedes Flüstern. In der Cafeteria muss jemand schon schreien, damit das auffällt. Und wenn Sie über Ihre Haut streichen, erscheint diese glatt und gleichmäßig. Schauen Sie aber genau hin, ist die Oberfläche durchaus nicht völlig glatt, aber die Abweichungen sind so klein, dass ihr Tastsinn das nicht bemerkt, es unterhalb der Wahrnehmungsschwelle liegt.

Die Unterschiedsschwelle hängt also von der Reizgröße ab. Man bezeichnet die Unterschiedsschwelle auch als „ebenmerklichen Unterschied“. Je größer ein Reiz ist, umso mehr muss sich ein zweiter Reiz davon unterscheiden, damit dieser zweite Reiz „ebenmerklich“ größer erscheint. Unsicherheitsintervall und Unterschiedsschwelle nehmen also mit der Reizgröße zu.

Überprüfen Sie nun, ob dies auch für Ihren Versuch gilt! Berechnen Sie dazu die Werte wie es hier beschrieben wurde (vielleicht haben Sie das ja schon beim Lesen direkt gemacht). Erstellen Sie damit eine Grafik wie in Abbildung 13 dargestellt. Sie finden dazu in Excel eine Vorlage, die Sie nutzen können. Im nächsten Abschnitt wird kurz beschrieben, wie man in Excel gute Grafiken erstellt. Wenn Sie das schon wissen, können Sie diesen Abschnitt überspringen.

3.4.4. Erstellung von guten Grafiken in Excel

Was ist eine gute Grafik? Man soll das Ergebnis, also die Datenpunkte gut und deutlich erkennen können und verstehen, was dargestellt ist. Das klingt so einfach, scheint aber erstaunlich schwierig zu sein, wenn man sieht, wie viele schlechte Grafiken es gibt. Und es sind doch einige Punkte zu beachten, um Daten gut darzustellen.

Um Datenpunkte gut und deutlich darzustellen, ist Folgendes zu beachten:

- Weiße Fläche (Hintergrund) des Diagramms: Kein Grau, keine Farbe
- Kein Gitternetz: Weder waagrecht noch senkrecht
- Datenpunkte groß genug: Symbole verwenden
- Farbe schwarz oder dunkle, gut erkennbare Farben
- Linien dick genug
- Wenn unterschiedliche Datenreihen: Unterschiedliche Symbole, Farben, Linienarten

Damit der Betrachter versteht, was dargestellt ist, gilt Folgendes:

- X- und y-Achse beschriften: Titel, Maßzahlen, Einheit
- Unterschiedliche Datenreihen in der Legende beschriften
- Ausführliche Bildunterschrift: Kurz beschreiben, was dargestellt ist

Und jetzt ganz konkret zu Ihren Daten. Markieren Sie den Datenbereich, der ungefähr so aussehen sollte wie in Tabelle 8. Am besten kopieren Sie dafür die entsprechenden Zeilen aus Ihrer Tabelle zusammen.

Tabelle 8: Der Datenbereich, der die Basis für die Grafik darstellt.

Streckenlänge	293	263	230	200	170	143	110
US	6	6	1	3	2	3	2

Wählen Sie dann ein X-Y-Diagramm aus (z.B. „Einfügen / Diagramme / Punkt XY“). Je nach Ihren Voreinstellungen sieht dies etwas unterschiedlich ist, entspricht aber in jedem Fall nicht den oben dargestellten Regeln. Und hier die Kurzanleitung:

- Größe einstellen: Bei „Diagrammtools / Format“ findet man Höhe und Breite. Stellen Sie die Höhe auf 5.7 cm, die Breite auf 8 cm ein. Dies passt gut für Text und Präsentationen.
- Schrift einstellen: Markieren Sie das ganze Diagramm und stellen Sie die Schriftart auf Arial, Schriftgröße 8 (eine klare Schrift für Zahlen, die Schriftgröße noch gut lesbar).
- Diagrammbereich einstellen (wenn Sie das Diagramm markiert haben, sollten rechts entsprechende Optionen auftauchen): Stellen Sie beim Rahmen „keinen Rahmen“ ein, bei der Füllung entweder „weiß“ oder „keine Füllung“.
- Zeichnungsfläche einstellen (markieren Sie den inneren Kasten im Diagramm): Stellen Sie die Füllung auf „weiß“ oder „keine Füllung“. Beim Rahmen stellen Sie „einfarbige Linie“, „schwarz“, „Stärke 1 Punkt“ ein.
- Achsenelemente einstellen: Bei „Diagrammtools / Layout / Beschriftungen“ finden Sie den Punkt „Diagrammelement hinzufügen“. Gehen Sie auf „Diagrammtitel“ und wählen Sie „keinen Titel“ (meist benötigt man keinen). Bei „Achsentitel“ wählen Sie „primär horizontal“ und „primär vertikal“ aus. Wenn die Titel erscheinen, können Sie gleich eine sinnvolle Beschriftung einfügen. Wenn „Gitternetzlinien“ ausgewählt sind, gehen Sie nochmal auf den Punkt und demarkieren Sie ihn.

- Achsenbeschriftungen einstellen: Wählen Sie zunächst die x-Achse aus, indem Sie auf die Zahlen an der Achse klicken. Stellen Sie hier die Linie wieder auf „einfarbig“ und „1 Punkt“. Bei „Achsoptionen“ stellen Sie bei „Teilstriche“ den „Haupttyp“ auf „Außen“. Bei dem Punkt „Zahl“ können Sie das Zahlenformat einstellen, wenn nötig. Wiederholen Sie dies für die y-Achse.
- Achseneinheit einstellen: Prüfen Sie, ob die Einheit der Achse (Anfang, Ende, Unterteilung) sinnvoll ist, um die Daten darzustellen. Es macht häufig wenig Sinn, die Achse bei „0“ anfangen zu lassen, wenn die Daten selbst relativ große Werte sind.
- Datenpunkte einstellen: Markieren Sie ihre Datenpunkte. Bei „Füllung und Linie“ stellen Sie bei „Linie“ die Linienfarbe und Stärke ein. Bei „Markierung“ finden Sie bei „Markierungsoptionen“ verschiedene Symbole und deren Farbe und können dies mit „Füllung“ und „Farbe“ weiter verändern.

Sie müssten auf diese Weise eine Grafik erhalten, die so ähnlich aussieht wie Abbildung 13. Sie finden dazu in Excel auch eine Vorlage. Sie können auch bei sich eine Trendlinie einfügen (Datenpunkte markieren, Rechtsklick, Trendlinie einfügen). Sie können diese Grafik jetzt markieren, kopieren und in Word oder Powerpoint einfügen. Ich würde empfehlen, dies über „Inhalte einfügen“, „Erweiterte Metadatei“ zu tun. Dann können Sie zwar in Word oder Powerpoint nicht mehr auf die Datenpunkte zugreifen, aber dafür verändert sich auch ihre Formatierung nicht. In Powerpoint können Sie die Grafik auf Foliengröße vergrößern, dann kann man das auch gut erkennen.

3.4.5. Das Weber'sche Gesetz

Abbildung 14 zeigt das Ergebnis dieses Versuchs für 61 studentische Probanden. Dabei wurde für jeden Probanden die Unterschiedsschwelle berechnet und dann über die 61 Probanden der Mittelwert gebildet. Auf der x-Achse ist die jeweilige Linienlänge in Pixeln dargestellt. Man sieht hier sehr deutlich, dass zwischen der Länge der Linie und der Unterschiedsschwelle ein linearer Zusammenhang besteht. Die Mittelwerte liegen fast exakt auf der Linie.

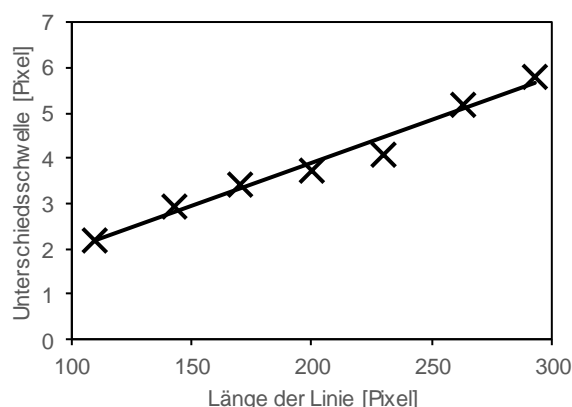


Abbildung 14: Ergebnis aus einer Stichprobe von $n = 61$ Studenten. Dargestellt ist der Mittelwert der Unterschiedsschwellen in Pixeln in Abhängigkeit von der „wahren“ Länge, außerdem die Regressionsgerade.

Was bedeutet ein solcher linearer Zusammenhang? Wenn man von der kürzesten Linie mit 110 Pixeln zu der zweiten mit 143 Pixeln geht, wird die Unterschiedsschwelle ungefähr 1 Pixel größer. Gleiches gilt, wenn man von der Linie mit 263 Pixeln zur längsten mit 293 Pixeln geht. Auch hier verändert sich die Unterschiedsschwelle (wenn man nach der Linie geht) um ungefähr einen Pixel. Auch die Verhältnisse müssen sich entsprechen: Wenn die Unterschiedsschwelle bei einer Linie von 110 Pixeln 2.2 Pixel ist, muss sie bei 220 Pixeln 4.4 Pixel sein.

Wie man in Abbildung 15 erkennt, ist dies auch tatsächlich ungefähr der Fall. Wenn man also diese Beziehung kennt, kann man damit vorhersagen, wie groß bei einer bestimmten Strecke die Unterschiedsschwelle sein wird. Man kann dies auch als Formel ausdrücken:

$$\frac{\Delta S}{S} = k$$

Dabei ist ΔS der Reizzuwachs. „ Δ “ bedeutet „Delta“, womit immer eine Veränderung gemeint ist. Der Reizzuwachs, der bei einem bestimmten Reiz S benötigt wird, um einen Unterschied zu bemerken, hängt linear von der Reizgröße ab. Deshalb ist der Bruch aus ΔS und S eine Konstante.

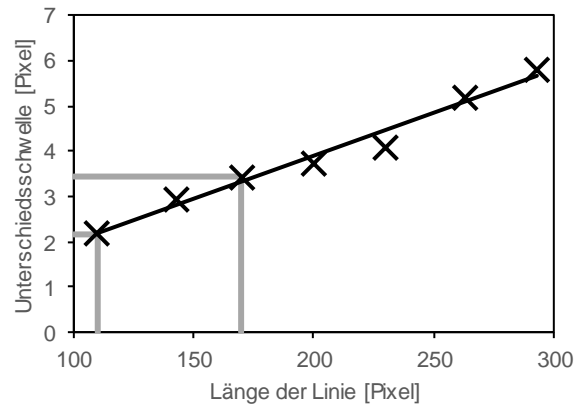


Abbildung 15: Dieselben Daten wie in Abbildung 14. Eingezeichnet sind die Unterschiedsschwellen bei Strecken von 110 und 170 Pixeln.

Bei 110 Pixeln ist die Unterschiedsschwelle 2.2 Pixel, bei 170 Pixeln beträgt sie 3.4 Pixel. In der Formel:

$$\frac{2.2}{110} = \frac{3.4}{170} = 0.02$$

Dass das auch für unseren Versuch gilt, zeigt Abbildung 16. Hier wurde für jede Linie der Bruch aus Unterschiedsschwelle und Linienlänge gebildet. Wie man erkennt, liegen die Werte nahezu auf einer Geraden, die parallel zur x-Achse ist. Das heißt, die Werte sind konstant. Dieser Bruch wurde 1834 von E.H. Weber aufgestellt und wird deshalb auch als „Weber-Bruch“ und die Konstante als „Weber-Konstante“ bezeichnet.

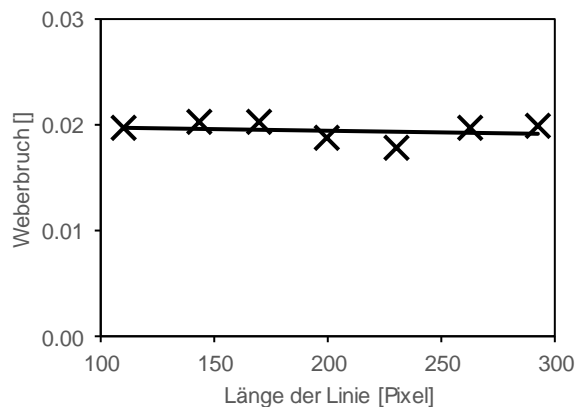


Abbildung 16: Dargestellt ist die Unterschiedsschwelle geteilt durch die Länge der Linie für die verschiedenen Linien des Versuchs für die n = 61 studentischen Probanden.

Diese Konstante ist bei unterschiedlichen Eigenschaften und unterschiedlichen Sinnen recht unterschiedlich. Je kleiner der Bruch ist, umso feinere Unterschiede kann man wahrnehmen. Je größer der Bruch ist, umso unempfindlicher für Unterschiede ist unser entsprechendes Sinnes-system.

Sie können jetzt den Weber-Bruch für Ihre eigenen Daten berechnen. Dazu müssen Sie für jede Strecke Ihre Unterschiedsschwelle durch die Länge der Strecke teilen. Dies ist in Tabelle 9 für das Beispiel dargestellt. Für die Darstellung wurde die Genauigkeit mit zwei Nachkommastellen in Excel eingestellt. Man erkennt, dass auch bei diesen einzelnen Daten das Weber'sche Gesetz annähernd gilt. Die Konstante liegt hier zwischen 0.01 und 0.02. Der Proband war also ähnlich unterscheidungsfähig wie die in Abbildung 16 dargestellte Gruppe von Studenten.

Tabelle 9: Berechnung des Weber-Bruchs für die eigenen Daten am Beispiel des einzelnen Probanden von oben.

Streckenlänge	293	263	230	200	170	143	110
US	6	6	1	3	2	3	2
US/Streckenlänge $\Delta S / S$	0.02	0.02	0.01	0.02	0.01	0.02	0.02

Wenn Sie bei einzelnen Strecken Abweichungen haben, dann kontrollieren Sie noch einmal Ihre Werte. Es kann immer sein, dass Sie bei einer Strecke aus Versehen zu früh gedrückt haben und die Schätzung dann sehr stark abwich. Oder Sie waren kurz unkonzentriert und haben sich ablenken lassen. Weil man derartige Fehler kaum vermeiden kann, wiederholt man derartige Versuche gerne häufig und mit vielen Probanden, so dass sich die Fehler dann „herausmitteln“.

Der Weber-Bruch ist sehr empfindlich für derartige Störungen. Man unterteilt diese Störungen in exogene (äußere) und endogene (innere). Wenn sich die Beleuchtung im Zimmer während des Versuchs ändert, wäre das ein exogener Störfaktor. Wenn Sie müde oder unkonzentriert werden, handelt es sich um einen endogenen Störfaktor. Um gute Ergebnisse zu erhalten, versucht man, solche Störvariablen so gut wie möglich auszuschalten. Deshalb hatten wir in der Instruktion bereits darauf hingewiesen, solche Störungen möglichst zu vermeiden. Wir hatten auch schon beim ersten Versuch diskutiert, dass sich die Versuchsergebnisse komplett ändern können, wenn Störvariablen systematisch wirken. Dort waren Wert und Größe konfundiert gewesen. Als wir im Entscheidungsexperiment die Störvariable „Wert“ ausgeschlossen haben, war unser Effekt nicht mehr zu finden. Es hatte also gar nicht die Größe gewirkt, sondern der Wert der Münzen. Wenn Sie bei dem Versuch jetzt mit einem dunklen Bildschirm gearbeitet haben, liegt Ihre schlechte Unterscheidungsfähigkeit und Ihr großer Weber-Bruch möglicherweise an der schlechten Beleuchtung und nicht daran, dass Ihre Sinne so unempfindlich sind. Die Eliminierung von Störvariablen ist also gerade für solche Wahrnehmungsexperimente außerordentlich wichtig.

3.4.6. Zusammenfassung

Wir hatten das Kapitel mit der Frage begonnen, ob das subjektive Erleben überhaupt messbar ist. Um diese Frage zu beantworten, mussten wir definieren, was unter Messung oder Metrik verstanden wird. Dabei haben wir die zwei wesentlichen Elemente einer Metrik kennengelernt:

- Die Qualität oder die Eigenschaft oder Dimension, die gemessen werden soll
- Die Quantität oder Messgenauigkeit, also die Frage nach der Unterscheidungsfähigkeit

Auf den ersten Aspekt waren wir nicht weiter eingegangen, sondern hatten das Erleben von Länge als Qualität für den Versuch ausgewählt. Dagegen war das zentrale Thema die Unterscheidungsfähigkeit: Wie fein kann man das Erleben der Länge messen? Wir hatten gesehen, dass diese Frage zu beantworten ist, indem man sich mit dem Gleich-Urteil näher beschäftigt. Die Aufgabe im Versuch war, eine Strecke zu mitteln, also in zwei gleich lange Strecken zu unterteilen. Die spezielle Frage war dann, wie gut und insbesondere wie genau dies gelingt.

Das erste zentrale Ergebnis des Versuchs war, dass es nicht einen Punkt gibt, auf den das „Gleich“-Erleben beschränkt ist, sondern dass es sich um einen ganzen Bereich handelt. Auch hier fanden wir also, dass physikalisch unterschiedliche Objekte als „gleich“ erlebt werden, also physikalische Welt und Erleben nicht übereinstimmen. Wir nannten diesen Bereich „Unsicherheitsintervall“. Wenn man fragt, ab wann das Gleich-Erleben in ein Ungleich-Erleben übergeht, kann man sowohl hin zu längeren als auch kürzeren Strecken untersuchen. Diesen Bereich des Übergangs bezeichneten wir als Unterschiedsschwelle, die dem halben Unsicherheitsintervall entspricht (oder umgekehrt: die Schwelle nach oben und unten zusammen ergibt das Unsicherheitsintervall).

Zur Berechnung haben Sie als statistische Kennwerte Varianz und Standardabweichung kennengelernt. Mit Hilfe der Standardabweichung wurde die Unterschiedsschwelle berechnet.

Mit diesem Kennwert ergab sich das zweite wichtige Ergebnis: Die Unterschiedsschwelle hängt ab von der Länge der Strecke. Dies wurde als Weber'sches Gesetz formuliert: Je größer der Reiz, desto größer ist auch die Unterschiedsschwelle. Dabei ist dieser Zusammenhang linear, sodass für eine bestimmte Eigenschaft oder Qualität die Weber-Konstante zu berechnen ist als Bruch aus Unterschiedsschwelle geteilt durch die Reizgröße. Damit lässt sich die Unterscheidungsfähigkeit des menschlichen Erlebens für bestimmte Eigenschaften als Zahlenwert angeben. Große Werte sprechen für eine schlechte Unterscheidungsfähigkeit, kleine für eine gute. Und weil dies vermutlich das erste psychologische Gesetz ist, das Sie kennenlernen, hier noch einmal die Formel:

$$\frac{\Delta S}{S} = k$$

Man muss sich die Bedeutung noch einmal klarmachen: Ein bestimmter Unterschied zwischen zwei Reizen wird nicht immer gleich wahrgenommen, sondern hängt von der Reizgröße ab. Zwei kleine Reize, die sich ein wenig unterscheiden, werden ganz deutlich als „unterschiedlich“ erlebt. Bei zwei großen Reizen, die sich um denselben Betrag unterscheiden, merkt man das dagegen nicht. Unser Erleben arbeitet also nicht additiv und bildet keine Differenzen, sondern Verhältnisse. Ebenmerklich mehr ist bei kleinen Reizen viel weniger als bei großen Reizen. Oder umgekehrt: Unser Erleben ist relativ. In einem bestimmten Kontext kann ein Reiz sehr stark erlebt, in einem anderen dagegen übersehen werden. Die Angst vor dem Zahnarzt hält einen davon ab, zur jährlichen Kontrolle zu gehen. Hat man dagegen Zahnschmerzen, fällt die Angst vor dem Hintergrund der Schmerzen so wenig ins Gewicht, dass man gerne den Zahnarzt aufsucht.

Im ersten Versuch hatten Sie gelernt, dass zwei gleiche Objekte als unterschiedlich beurteilt werden. Sie hatten gesehen, dass neben der Größe der Wert eine Rolle spielt. Menschen isolieren offensichtlich nicht die Größe vom Wert, sondern für Menschen gehört der Wert zur Größe dazu. So gesehen wurde in diesem Versuch gezeigt, dass die psychische Qualität oder Eigenschaft nicht mit der physikalischen übereinstimmen muss. Bei dem zweiten Versuch ging es um die Quantität, also die Genauigkeit der Wahrnehmung. Auch hier wurde gezeigt, dass das Erleben systematisch von der physikalischen Messung abweicht. Diese Abweichung konnte sogar in einem Gesetz gefasst werden: Das Erleben berücksichtigt nicht nur Unterschiede, sondern setzt diese ins Verhältnis zur Reizgröße. Das kann durchaus positiv gesehen werden, da beim Erleben damit immer das Umfeld, der Kontext, berücksichtigt wird. Diese Anpassungsfähigkeit hilft dabei, in unterschiedlicher Umwelt gut wahrnehmen zu können. Aber diese Überlegungen stehen nicht im Zentrum dieses Buches.

4. Versuch Konstanzverfahren: Das psychische Maß – welche Dimensionen gibt es?

4.1. Die zwei Aspekte einer Metrik

Im letzten Kapitel hatten wir dargestellt, dass jede Metrik oder Messung zwei Aspekte unterscheidet, die Frage nach den Eigenschaften, die gemessen werden sollen, und nach der Messgenauigkeit. In diesem Kapitel hatten wir uns mit der Messgenauigkeit beschäftigt, die im Bereich des Erlebens mit der Unterscheidungsfähigkeit von Reizausprägungen beschrieben wurde. Das zentrale Ergebnis war, dass die Unterscheidungsfähigkeit von der Reizausprägung abhängt: Je größer ein Reiz ist, umso größer muss auch der Unterschied zu einem zweiten Reiz sein, damit dieser als unterschiedlich wahrgenommen wird. Daraus war das Weber'sche Gesetz abgeleitet worden: Reize müssen um den gleichen Prozentsatz größer werden, damit man diese unterscheiden kann.

- Die Metrik des Psychischen ist prozentual.

In der Physik ist das anders: Dort ist z.B. ein Unterschied von 1 mm Länge immer derselbe Unterschied, egal, ob die Körper insgesamt 1 m lang sind oder 1 cm. Oder: 1 mm länger bleibt immer 1 mm länger.

- Die Metrik der Physik ist additiv.

Damit unterscheidet sich die Metrik der Physik und des Psychischen in Bezug auf die Unterscheidungsfähigkeit oder Messgenauigkeit. Wie sieht es jetzt mit dem zweiten Aspekt, der Dimension, aus?

Die klassische Physik kennt nur drei Dimensionen: Länge, Gewicht und Zeit. Andere Wissenschaften und das menschliche Erleben sind wesentlich vielfältiger. Aber welche Dimensionen sind beim Erleben relevant?

4.2. Die psychischen Dimensionen

Als Sie beschlossen haben, Psychologie zu studieren, hatten Sie sich sicherlich nicht vorgestellt, etwas über das menschliche Erleben beim Einschätzen von Größe zu lernen. Wenn man an das menschliche Erleben denkt, fallen einem spontan ganz andere Begriffe ein: Freude, traurig sein, Müdigkeit, Liebe, Langeweile usw. Diese Liste ist sehr lang und tatsächlich beschäftigt die Frage, welches die menschlichen Erlebensweisen sind und wie sie entstehen, bis heute die Psychologie. Auch dieses Buch kann darauf keine vollständige Antwort geben. Hier soll es viel mehr darum gehen, wie man einzelne Erlebensweisen wissenschaftlich beschreiben kann – eben die Frage nach der Metrik des Psychischen beantworten. Dies werden wir anhand unterschiedlicher Erlebensweisen oder Dimensionen untersuchen. Aber als kleiner Exkurs sollen zumindest einige Bemerkungen zu den Dimensionen des Erlebens folgen.

Wie schon bei der Frage der Messgenauigkeit kann auch bei der Frage nach den Dimensionen der Blick in andere Wissenschaften einen Hinweis darauf geben, welche Dimensionen für das Erleben relevant sind. In der Chemie werden z.B. Eigenschaften von Stoffen beschrieben, indem auf Wahrnehmungen zurückgegriffen wird:

- Reine Fette sind geschmack- und geruchlos und fühlen sich fettig an.

Ähnliches findet sich in der Biologie:

- Der Hut des Fliegenpilzes ist rot mit weißen, runden Flecken.

In beiden Wissenschaften werden Eigenschaften von Stoffen, von Pflanzen, von Tieren beschrieben, indem bestimmte Wahrnehmungen aufgezählt werden. Dies legt die Idee nahe, dass

die Dimensionen unseres Erlebens mit unserer Wahrnehmung zusammenhängen. Diese wiederum ist durch unsere Sinne bestimmt. Die wichtigsten Sinne sind Sehen, Hören, Schmecken, Fühlen und Riechen. Die unterschiedliche Wahrnehmung wird durch verschiedene Sinnesorgane ermöglicht (z.B. Auge, Ohr, Zunge, Haut, Nase), in denen jeweils Rezeptoren enthalten sind, die die Reize in der Umwelt für den Menschen wahrnehmbar machen. Grundsätzlich gilt:

- Ein Reiz trifft auf einen Rezeptor, der dann Informationen über den Reiz über Nervenbahnen ins Gehirn, das zentrale Nervensystem, leitet.

Erleben bedeutet ja zunächst, dass bestimmte Reize oder Objekte der Umwelt in uns mental repräsentiert sind. Damit ist das Erleben weitgehend durch die Wahrnehmung bestimmt. Wenn also die menschliche Wahrnehmung dadurch entsteht, dass unterschiedliche Rezeptoren Informationen liefern, könnte die Untersuchung dieser verschiedenen Rezeptoren zu den psychischen Dimensionen führen: Als Mensch kann man das erleben, wofür es entsprechende Rezeptoren gibt, bzw. worauf diese reagieren. Am Beispiel des Sehens: Die Sehrezeptoren im Auge reagieren auf Helligkeit, was sicherlich eine psychische Dimension darstellt. Wir können hell und dunkel unterscheiden und verschiedene Stufen der Helligkeit beschreiben. Weiter gibt es unterschiedliche Sehrezeptoren, die auf unterschiedliche Wellenlängen von Licht reagieren, was vom Mensch als unterschiedliche Farben erlebt wird. Weiter reagieren unterschiedliche Rezeptoren, je nachdem, wo sich bestimmte Objekte in der Umwelt befinden. Damit erleben wir ein „Oben“ und „Unten“, „Rechts“ und „Links“. Aber dann wird es schwierig: Warum ist ein Bild schön? Gibt es einen „Schönheitsrezeptor“? Wann wird eine Farbe „grell“? Man konnte bislang keinen eigenen Rezeptor für „Grellheit“ finden, sondern es handelt sich um eine komplexe Kombination aus Helligkeit und Farbigkeit, die zu diesem Erleben führt.

Gleiches gilt für die anderen Sinne: Je nach Lautheit eines Tones reagieren die Hörrezeptoren unterschiedlich stark. Die Tonhöhe oder Frequenz des Tones wird mit Hilfe unterschiedlicher Rezeptoren für hohe und tiefe Töne wahrgenommen. Der Herkunftsort eines Geräuschs wird wahrgenommen, weil die Rezeptoren im linken und rechten Ohr zeitlich leicht unterschiedlich reagieren. Aber woher wissen wir, ob gerade ein Mann oder eine Frau spricht? Warum ist eine Stimme angenehm, eine andere rau oder schrill? Was unterscheidet eine Geige von einer Trompete? Auch hier gilt: Es gibt eine Vielzahl von akustischen Eigenschaften, die man als Mensch unmittelbar wahrnimmt, für die sich aber nicht jeweils eigene Rezeptoren nachweisen lassen, sondern deren Basis ein komplexes Muster der Erregung in wenigen Rezeptoren darstellt. Umgekehrt bedeutet dies aber auch, dass die Analyse der physiologischen Grundlagen, die Identifikation der Rezeptoren, auf denen unsere Wahrnehmung beruht, zu kurz greift. Das menschliche Erleben ist wesentlich vielfältiger als die wenigen Arten von Rezeptoren, die man physiologisch unterscheiden kann.

- Die psychischen Dimensionen gehen weit über einfache Sinnesqualitäten hinaus.
- Die Psychologie muss die Dimensionen ihrer Metrik selbst bestimmen.

Auch das erste Experiment unterstützt diesen Gedanken: Die Größe einer Münze ist nicht nur von der physikalisch messbaren Größe abhängig, sondern auch von ihrem Wert. Dieses Erleben des Werts beruht auf der Erkennung bestimmter Eigenschaften der Münze, vor allem natürlich auf den Schriftzeichen, die aufgeprägt sind. Die Wahrnehmung der Schriftzeichen führt aber erst dadurch zum Erleben eines bestimmten Werts, weil wir gelernt haben, dass man mit den Münzen unterschiedlich wertvolle Gegenstände kaufen kann. Das Gelernte, die Vorerfahrung bestimmen damit ganz wesentlich das Erleben. Die Wahrnehmung allein kann das Erleben nicht erklären.

Damit wird aber auch deutlich, dass man die Psychologie als eigene Wissenschaft braucht: Physik und Physiologie reichen eben nicht aus, das Erleben von Menschen zu erklären. Wir

hatten das unter dem Begriff der Metrik des Psychischen schon angesprochen. Bei dem zweiten Experiment war deutlich geworden, dass sich die Metrik des Psychischen hinsichtlich der Messgenauigkeit von der Physik unterscheidet und eine prozentuale Metrik ist. Die Diskussion hier zeigte, dass auch bei dem zweiten Aspekt der Metrik, der Frage nach den Dimensionen, die Psychologie eigenständig arbeiten muss. Diese Frage ist nur zu erklären, indem man im Erleben der Menschen nach unterscheidbaren Eigenschaften sucht.

Welche Erlebensbereiche hat der Mensch? Sie sehen, dass diese Frage nicht einfach zu beantworten ist. Wir werden in diesem Buch versuchen, für einige sehr grundlegende Erlebensbereiche deren Metrik weiter zu beschreiben. Die ersten beiden Experimente kamen aus dem Bereich des visuellen Erlebens. Beim nächsten Versuch wechseln wir deshalb zum Hören, also zum akustischen Erleben.

4.3. Anleitung zum Versuch Konstanzverfahren

Im Versuch Streckenmitteln haben Sie ein psychisches Maß (Erleben der Streckenlänge) kennengelernt. Mit einer bestimmten Methode hatten wir die Messgenauigkeit bestimmt und konnten zeigen, dass wir die Länge sehr gut einschätzen können, wobei die Genauigkeit entsprechend dem Weber'schen Gesetz von der Länge der Striche abhing. Denken Sie noch einmal kurz an die Methode zurück! In dem Versuch sollten Sie einen Strich in zwei gleichlange Striche aufteilen. Hier verbarg sich die Längenschätzung: „Gleich lang“ bedeutet, dass der linke Strich genau dieselbe Länge aufweist wie der rechte Strich. Die Abweichung von der tatsächlichen halben Strichlänge diente uns als Grundlage für die Berechnung der Genauigkeit der Längenschätzung. Im ersten Versuch waren wir beim Größenschätzen etwas anders vorgegangen: Man sollte sich eine bestimmte Münze vorstellen und dann die Größe des Kreises so einstellen, dass sie der Vorstellung von der Münze entsprach. Sie kennen damit bereits zwei Arten, mit denen man die Metrik des Psychischen untersuchen kann.

Allerdings war diesen beiden Arten eines gemeinsam: Sie haben nach einem bestimmten Reiz gesucht, indem Sie den Reiz solange verändert haben, bis er Ihrer Vorstellung entsprach (bei den Münzen) oder Sie keinen Unterschied mehr gesehen haben (der der Strichteilung). Bei beiden Verfahren waren Sie ganz ähnlich vorgegangen: Sie hatten mit einem zu kleinen oder zu großen Reiz begonnen und sich dann dem richtigen Reiz angenähert, zunächst in großen Schritten, dann in kleineren, wobei Sie um den richtigen Reiz gependelt sind, bis sie sich schließlich entschieden haben. Man nennt dieses Verfahren ein Herstellungsverfahren.

- Bei dem Herstellungsverfahren verändert der Proband selbst den Reiz aktiv so lange, bis er seinen Anforderungen entspricht.

Beim folgenden Versuch werden Sie eine zweite Methode kennenlernen. Wir werden später ausführlich diskutieren, welche Vor- und Nachteile verschiedene Methoden haben. Aber eines vorweg: Je nach Erlebensbereich können verschiedene Methoden besser oder schlechter geeignet sein. Für einen Psychologen ist es daher wichtig, unterschiedliche Methoden und ihren Anwendungsbereich zu kennen. Diese Methodenkompetenz ist eine ganz wichtige Eigenschaft eines erfolgreichen Psychologen.

Das Herstellungsverfahren hat einen wesentlichen Nachteil: Der Proband muss den Reiz selbst aktiv variieren können. Das ist nicht immer einfach möglich. Denken Sie z.B. an das Gewicht: Wie stellt man ein bestimmtes Gewicht stufenlos ein? Bei alten Wagen gab es bestimmte Gewichte, mit denen man eine Waagschale schwerer oder leichter machen konnte, aber eben nur mit einer bestimmten Genauigkeit, die von den Gewichten abhing. Wenn man die Genauigkeit des menschlichen Schwereerlebens untersuchen will, bräuchte man sehr feine, kleine Gewichte.

Hier könnte es sinnvoll sein, eine andere Methode zu verwenden. Im folgenden Versuch werden Sie das Konstanzverfahren als zweite Methode kennenlernen. Wie funktioniert das?

Stellen Sie sich vor, Sie möchten einen neuen Lippenstift kaufen. Bei einer Freundin hatten Sie eine Farbe gesehen, die Ihnen sehr gut gefallen hat. Im Geschäft lassen Sie sich unterschiedliche Lippenstifte vorlegen, solange, bis Sie einen gefunden haben, der von der Farbe her dem Ihrer Freundin entspricht. Bei diesem Vorgehen haben Sie einen Standardreiz (Farbe des Lippenstifts der Freundin) und beurteilen eine Reihe von Vergleichsreizen (die Lippenstifte im Geschäft). Ihr Urteil ist jeweils „gleich“ oder „nicht gleich“. Konstanzverfahren heißt diese Methode, weil der Standardreiz konstant bleibt und man Vergleichsreize mit diesem vergleicht. Man nennt dies auch eine Methode der Urteilsfindung, weil hier nach dem „Gleich“-Urteil gesucht wird. Das Herstellungsverfahren wird dagegen als Methode der Reizfindung bezeichnet, weil hier ein Reiz gesucht wird, der einer Vorstellung oder Wahrnehmung entspricht.

- Das Herstellungsverfahren ist eine Methode der Reizfindung. Der Proband variiert aktiv die relevante Reizeigenschaft, bis ein bestimmtes Kriterium erreicht ist.
- Das Konstanzverfahren ist eine Methode der Urteilsfindung. Ein Standardreiz wird vorgegeben und der Proband beurteilt eine Reihe von Vergleichsreizen daraufhin, ob sie dem Standardreiz in der relevanten Reizeigenschaft entsprechen oder nicht.

In dem Versuch werden wir auch ein anderes Sinnessystem untersuchen. Es geht um Tonhöhe. Auch hier stellt sich die Frage, wie die physikalischen Reize (hier: Töne unterschiedlicher Frequenz) erlebt werden (hier: wie hoch der Ton ist). Gilt auch für diese Dimension das Weber'sche Gesetz? Wie empfindlich ist dieses Erleben im Vergleich zur Längen- oder Größenswahrnehmung? Sie können diese Frage beantworten, indem Sie die Unterschiedsschwellen mit denen aus den ersten beiden Versuchen vergleichen.

4.4. Der Versuch Konstanzverfahren

In dem Versuch soll die Unterschiedsschwelle für Tonhöhe bestimmt werden. Als Methode wird das Konstanzverfahren eingesetzt. Sie hören also immer einen Standardreiz, in diesem Fall ein Sinus-Ton von 1000 Hz. Direkt danach wird ein zweiter Ton präsentiert, ein Vergleichsreiz. Sie sollen angeben, ob der zweite Ton

- gleich hoch (0),
- tiefer (-1), oder
- höher (1)

ist. Da es um die Unterschiedsschwelle geht, dürfen sich die Vergleichsreize nur sehr wenig von dem Standardreiz unterscheiden. Es soll ja genau der Bereich gefunden werden, in dem das Urteil unsicher ist. Das führt dazu, dass der Versuch von den meisten Probanden als sehr schwierig und anstrengend erlebt wird. Lassen Sie sich also nicht frustrieren! Die Töne hören sich tatsächlich alle sehr gleich an. Es liegt nicht daran, dass Sie Tonhöhen schlecht unterscheiden können, sondern daran, dass wir den Unsicherheitsbereich eingrenzen möchten.

Insgesamt gibt es zwölf Vergleichsreize. Einer davon ist genau gleich hoch wie der Standardreiz, sechs sind tiefer, fünf höher. Da wir uns hier im Bereich der Unsicherheit bewegen, ist es notwendig, die zwölf Vergleiche mehrfach zu wiederholen. Dabei ist die Reihenfolge der Reize immer zufällig. Das soll Erwartungen auf Ihrer Seite vorbeugen. Es hat sich gezeigt, dass 10 Wiederholungen gut geeignet sind, um eine individuelle Schwelle zu bestimmen. Sie müssen also 120 Vergleiche durchführen, bei denen Sie sich sehr stark konzentrieren müssen. Von da her macht es sicherlich Sinn, zwischendurch Pausen zu machen.

Ganz wichtig ist, dass Sie bei dem Versuch ungestört sind. Machen Sie den Versuch in einem Raum, in dem möglichst keine Außengeräusche störend zu hören sind. Oder setzen Sie Kopfhörer auf.

Wenn Sie alles soweit verstanden haben, können Sie das Programm aufrufen und den Menüpunkt „Konstanzverfahren“ auswählen. Sie erhalten hier noch einmal alle Informationen für den Versuch. Sie sollten hier auch das Lautsprechersymbol anwählen und den Standardreiz anhören. Stellen Sie die Lautstärke des Computers so ein, dass Sie diesen Ton gut hören können, er weder zu laut noch zu leise ist. Dann kann es losgehen.

Wenn Sie den entsprechenden Knopf drücken, hören Sie kurz den Standardreiz, gefolgt von einer Pause und dem jeweiligen Vergleichsreiz. Danach erscheint eine Auswahlmöglichkeit, bei der Sie eingeben können, ob der Vergleichsreiz tiefer, gleich oder höher war.

Führen Sie jetzt bitte den Versuch durch!

Und hier geht es zum Versuch:

- <https://cgi-psych.rz.tu-bs.de>

4.5. Auswertung und Ergebnisse

Wie bei den ersten beiden Versuchen soll auch hier die Messgenauigkeit unseres Erlebens bestimmt werden. Kurz zur Erinnerung: Die Messgenauigkeit wurde beschrieben, indem die Streuung um den Punkt subjektiver Gleichheit bestimmt wurde. Dies ist in Abbildung 17 dargestellt. Die kleinen Linien sind die Ergebnisse aus den verschiedenen Durchgängen. Als Mittelwert aus diesen Messwerten hatten wir den Punkt subjektiver Gleichheit bestimmt. Um das Unsicherheitsintervall zu bestimmen, war die Standardabweichung als Streuungsmaß berechnet worden. In der Grafik ist diese je einmal nach links und rechts abgetragen worden. Dies entspricht dem Unsicherheitsintervall.

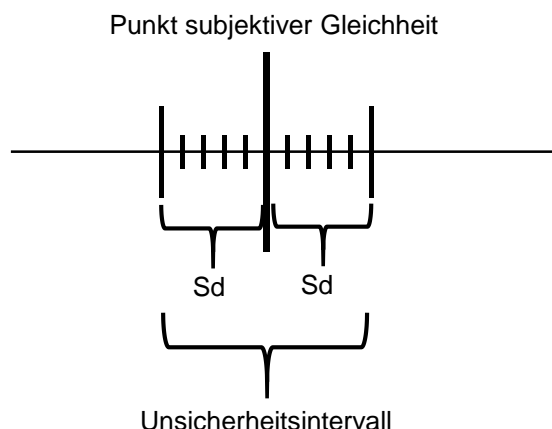


Abbildung 17: Unsicherheitsintervall und Punkt subjektiver Gleichheit. „Sd“ steht für Standardabweichung.

Diese Standardabweichung hatten wir auch als Maß für die Unterschiedsschwelle herangezogen. Die Unterschiedsschwelle folgt wiederum in einem Bereich mittlerer Reizgrößen dem Weber’schen Gesetz:

$$\frac{\Delta S}{S} = k$$

Die Konstante, die man so berechnen kann, kann als Maß für die Sensitivität oder Empfindlichkeit des jeweiligen Erlebensbereichs verwendet werden.

Von da her sollen diese Maße auch für das Konstanzverfahren berechnet werden. Wie berechnet man hier das Unsicherheitsintervall? Abbildung 18 veranschaulicht die Situation beim Konstanzverfahren. Bei sehr tiefen und sehr hohen Tönen ist das Urteil ganz sicher. Je näher man sich von der Tonhöhe her an den Standardreiz annähert, umso häufiger kann man sich nicht sicher entscheiden. Irgendwo in dem Bereich zwischen dem sicheren Urteil und diesem mittleren Bereich müssen also die Grenzen des Unsicherheitsintervalls liegen. Wie legt man diese fest?

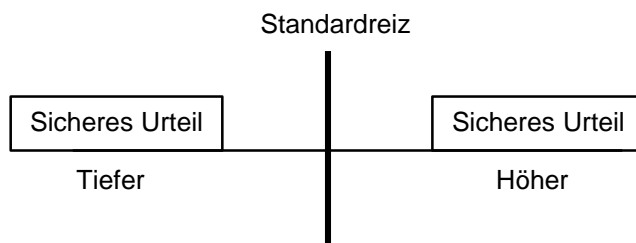


Abbildung 18: Die Urteilsituation beim Konstanzverfahren.

Ein völlig sicheres Urteil liegt dann vor, wenn bei einem Reiz alle Urteile eindeutig waren, also z.B. jedes Mal „tiefer“ angeklickt wurde. Umgekehrt ist ein Reiz wiederum ganz sicher nicht mehr „tiefer“, wenn dieses Urteil gar nicht mehr verwendet wird. Von da her macht es Sinn, als Grenze den Wert von 50% Urteilen zu verwenden, also den Reiz, bei dem der Proband fünf von zehn Mal „tiefer“ angeklickt hat. Abbildung 19 zeigt die Urteile eines Probanden. Auf der linken Seite erkennt man, dass bei den tiefsten Reizen das Urteil absolut sicher war, da alle 10 Urteile jeweils „tiefer“ waren. Ab 1004 Hz taucht kein Tiefer-Urteil mehr auf. In dem Bereich dazwischen liegt also die Grenze und beginnt das Unsicherheitsintervall. Rechts in Abbildung 19 ist der Bereich zwischen 995 und 1003 Hz vergrößert dargestellt. Bei 50% der Urteile, also 5 der 10 Urteile ist eine Linie eingezeichnet. Der Schnittpunkt mit der Funktion liegt bei 998.5 Hz, was auf diese Weise grafisch als Beginn des Unsicherheitsintervalls definiert werden kann.

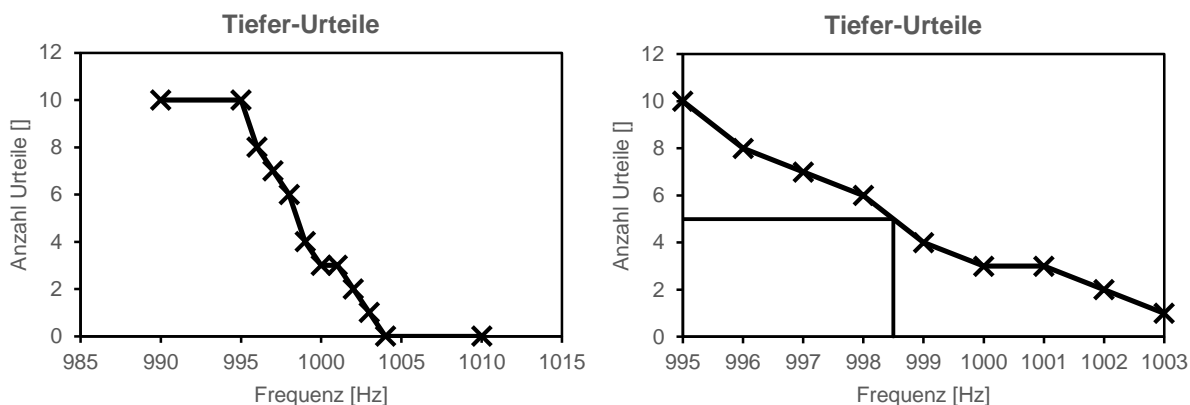


Abbildung 19: Häufigkeit der Tiefer-Urteile bei einem Probanden. Auf der rechten Seite ist der Bereich zwischen 995 Hz und 1003 Hz vergrößert dargestellt.

In gleicher Weise wird nun die andere Seite des Unsicherheitsintervalls bestimmt. Abbildung 20 zeigt die entsprechenden Werte des Probanden. Ab 1004 Hz sind die Urteile alle „höher“. Die Kurve fällt dann und bei 1001 Hz lauten 5 der 10 Urteile „höher“. Damit liegt die obere Grenze des Unsicherheitsintervalls bei 1001 Hz. Damit ist das Unsicherheitsintervall bei diesem Probanden wie folgt bestimmt:

- Unsicherheitsintervall UI = [998.5, 1001]

Das Unsicherheitsintervall ist also 2.5 Hz breit. Die Unterschiedsschwelle ist ja, wie oben dargestellt, die Hälfte des Unsicherheitsintervalls:

- Unterschiedsschwelle $\Delta S = 1.25 \text{ Hz}$

Damit kann die Weber-Konstante berechnet werden. Der Vergleichston hatte 1000 Hz. Die Unterschiedsschwelle ist 1.25 Hz. Also ist:

$$k = \frac{\Delta S}{S} = \frac{1.25 \text{ Hz}}{1000 \text{ Hz}} = 0.00125$$

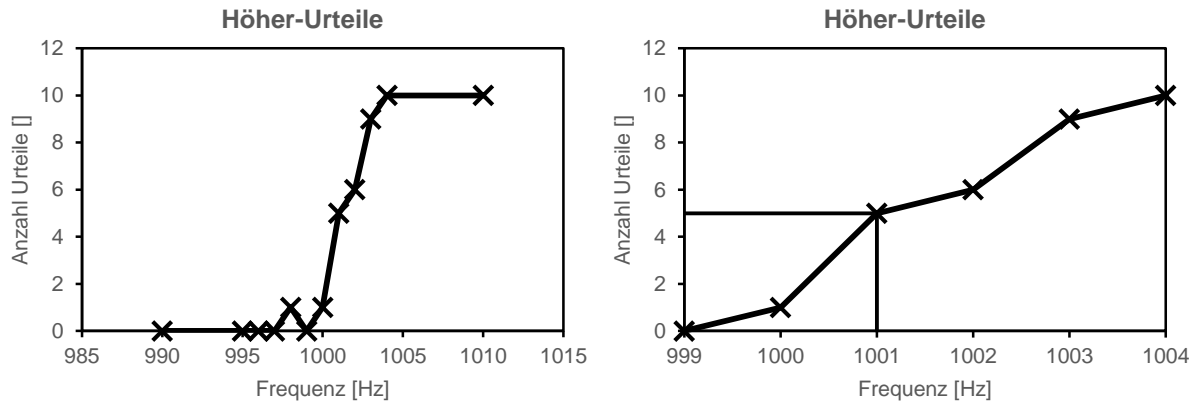


Abbildung 20: Häufigkeit der Höher-Urteile bei einem Probanden. Auch hier ist wieder rechts der relevante Bereich vergrößert dargestellt.

Beim Streckenmitteln (s. Abbildung 16 auf Seite 35) hatten wir im Beispiel eine Weber-Konstante von 0.02 erhalten. Die Konstante für die Beurteilung von Tonhöhen ist im Beispiel wesentlich kleiner. Das heißt, dass die Unterscheidungsfähigkeit für Tonhöhen wesentlich besser ausgeprägt ist als die für die Länge von Strecken. Angesichts der Leistungen, die unser auditives System zum Beispiel bei der Sprachwahrnehmung vollbringt, ist dies nicht verwunderlich. Dieses System kann außerordentlich gut verschiedene Frequenzen unterscheiden, was uns erst befähigt, Sprache zu verstehen.

Werten Sie jetzt Ihre eigenen Daten entsprechend aus. Zählen Sie dazu pro Reiz die Anzahl der „höher“- , „gleich“- und „tiefer“-Urteile aus und tragen Sie sie in die entsprechende Tabelle ein. Sie können dann in Ihren Abbildungen so wie hier dargestellt grafisch die Schwellen bestimmen. Mit einiger Wahrscheinlichkeit werden Ihre eigenen Daten nicht so stetig verlaufen wie die hier im Beispiel gezeigten. Abbildung 21 zeigt ein weiteres Beispiel.

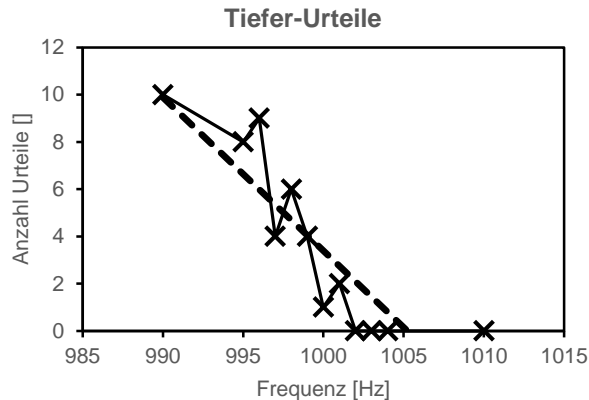


Abbildung 21: Beispiel für eine weitere Verteilung von Tiefer-Urteilen. Die gestrichelte Linie ist eine Trendlinie zur Annäherung der Daten.

Bei dieser Kurve ist es aus den Originaldaten schwierig, die 50%-Grenze zu bestimmen. Hier macht es Sinn, die Verteilung der Urteile durch eine Linie anzunähern. In Excel können Sie dazu eine Trendlinie einfügen (Diagrammtools – Entwurf – Diagrammelement hinzufügen – Trendlinie – Linear). Sie können dann den 50%-Wert dieser Trendlinie als Messwert nehmen. Zeichnen Sie dazu eine waagerechte Linie bei 5 Urteilen und lesen Sie ab, an welcher Stelle der x-Achse diese ihre Trendlinie schneidet. Wiederholen Sie dies für die „höher“-Urteile und bestimmten Sie so Ihr Unsicherheitsintervall. Berechnen Sie daraus dann die Unterschiedsschwelle und die Weber-Konstante. Tragen Sie die Werte in das Buch ein!

- Untere Grenze:
- Obere Grenze:
- Unsicherheitsintervall:
- Weber-Konstante:

Wir haben damit die Messgenauigkeit unseres akustischen Systems im Hinblick auf die Tonhöhe bestimmt. Außerdem ist ja interessant, wie sich die physikalischen Eigenschaften im Erleben abbilden. Dies wird mit dem Punkt subjektiver Gleichheit angesprochen: Bei welcher Tonhöhe erscheint uns der Vergleichston denn dem Standardton „gleich“? Wir hatten am Anfang des Kapitels gezeigt (s. Abbildung 17), dass der Punkt subjektiver Gleichheit in der Mitte des Unsicherheitsintervalls liegt. Wir können also umgekehrt das Unsicherheitsintervall halbieren, um so den Punkt subjektiver Gleichheit zu bestimmen. Dies ist in Abbildung 22 unten dargestellt. Das Unsicherheitsintervall in dem Beispiel reichte von 998.5 Hz bis 1001 Hz, war also 2.5 Hz breit. Halbiert man dies und addiert es zur unteren Grenze, so ergibt sich als Punkt subjektiver Gleichheit PSG:

$$PSG = \frac{2.5 \text{ Hz}}{2} + 998.5 \text{ Hz} = 1.25 \text{ Hz} + 998.5 \text{ Hz} = 999.75 \text{ Hz}$$

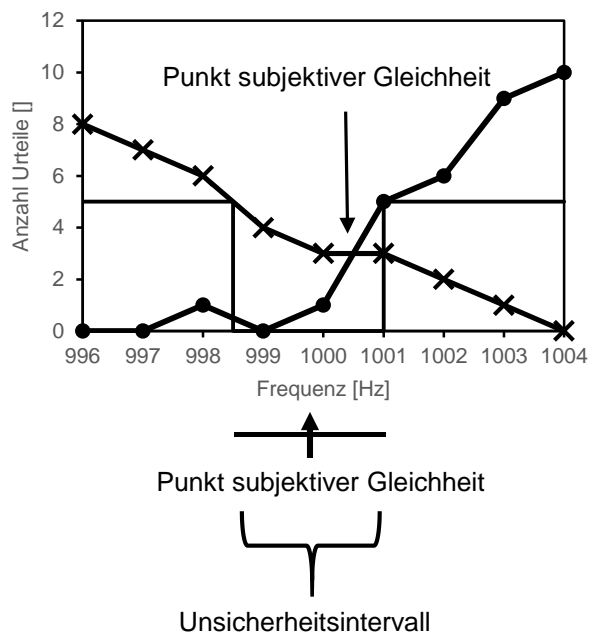


Abbildung 22: Bestimmung des Punkts subjektiver Gleichheit mit zwei Möglichkeiten. Zur Erklärung, s. Text.

Damit entspricht der Punkt subjektiver Gleichheit nicht ganz dem Standardton von 1000 Hz. Solche geringen Abweichungen können allerdings auch zufällig entstehen. In der Abbildung 22 ist noch eine weitere Möglichkeit dargestellt, den Punkt subjektiver Gleichheit zu bestimmen: Der Schnittpunkt der „tiefer“- und „höher“-Urteile entspricht ja ebenfalls einem Punkt, an dem man sich nicht entscheiden kann. Liest man diesen Wert aus der Abbildung ab, so ergibt sich als PSG = 1000.5 Hz, also ein etwas größerer Wert als der Standardton. Und schließlich haben wir außerdem noch die „gleich“-Urteile, die wir bislang noch nicht genauer untersucht haben. Mit diesen können wir ebenfalls einen PSG bestimmen als den Reiz, bei dem am meisten „gleich“-Urteile auftreten. Abbildung 23 zeigt die entsprechende Kurve. Bei 999 Hz und 1000 Hz sind die „gleich“-Urteile am häufigsten mit jeweils 6 Urteilen. Hier wäre der PSG demnach der Mittelwert aus diesen beiden Frequenzen, also 999.5 Hz.

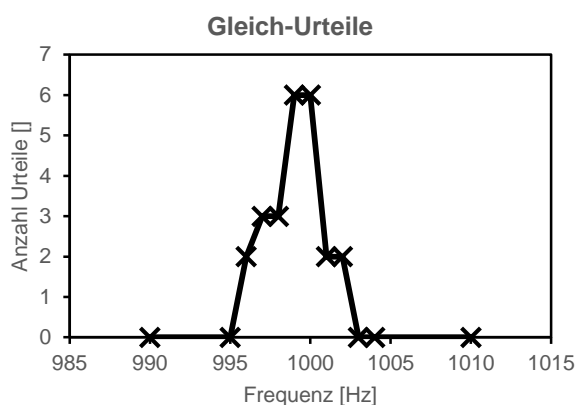


Abbildung 23: Häufigkeitsverteilung der „gleich“-Urteile.

Zusammenfassend ergeben sich damit beim Konstanzverfahren drei Möglichkeiten, den Punkt subjektiver Gleichheit zu bestimmen:

- Die Mitte des Unsicherheitsintervalls – hier 999.75 Hz
- Der Schnittpunkt der „tiefer“- und „höher“-Urteile – hier 1000.5 Hz

- Der Reiz mit den häufigsten „gleich“-Urteilen – hier 999.5 Hz

Sie sehen, dass die Ergebnisse nicht identisch sind. Je nach Methode gibt es leicht unterschiedliche Zahlen. Auf diesen Punkt werden wir später noch ausführlicher eingehen. Letztlich ist wichtig, dass bei der Darstellung von Ergebnissen immer die Methode beschrieben werden muss, um die Ergebnisse nachvollziehen und möglicherweise replizieren zu können. Jede der drei Methoden ist für sich sinnvoll und begründbar. Es gibt häufig in der Wissenschaft kein „richtig“ und „falsch“, sondern die Frage ist, welche Berechnung dem am besten entspricht, was man als Wissenschaftler beschreiben will, also hier den Punkt subjektiver Gleichheit. Welche der drei Definitionen oben entspricht am besten Ihrem Verständnis von „Gleichheit“? Sie können durchaus unterschiedlicher Meinung sein. Viele Diskussionen erübrigen sich allerdings, wenn man sich vorher darüber verständigt, was man eigentlich meint, wie man bestimmte Begriffe definiert. Häufig liegt hier die Ursache für die Diskussionen und gar nicht in der Sache an sich.

Doch zurück zur Methodik. Der Punkt subjektiver Gleichheit liegt in der Mitte des Unsicherheitsintervalls. Dessen Breite ergibt sich durch die Unterschiedsschwelle, die nach oben und unten abgetragen wird. Dabei gibt es beim Konstanzverfahren eine Besonderheit: Je häufiger ein Proband das „gleich“-Urteil verwendet, umso breiter wird das Unsicherheitsintervall, da die 50%-Grenze dann eher erreicht wird. Interessant ist, dass die Häufigkeit des „gleich“-Urteils auch davon abhängt, welche Instruktion man den Probanden gibt:

- Instruktion 1: „Verwenden Sie das Urteil ‚gleich‘ nur dann, wenn Sie wirklich trotz aller Konzentration keinen Unterschied feststellen können! Sie sollten möglichst selten ‚gleich‘ sagen, weil eigentlich alle Reize unterschiedlich sind.“
- Instruktion 2: „Verwenden Sie die Urteile ‚höher‘ oder ‚tiefer‘ nur, wenn Sie sich ganz sicher sind. Sie sollen auf keinen Fall raten. Es wirklich darum, deutliche Unterschiede zu finden. Scheuen Sie sich nicht ‚gleich‘ zu urteilen!“
- Instruktion 3: „Wenn der Vergleichston tiefer oder höher ist, dann wählen Sie dies entsprechend aus. Sagen Sie ‚gleich‘, wenn Sie keinen Unterschied hören.“

Sie könnten selbst relativ leicht mit einem kleinen Versuch herausfinden, welche Wirkungen diese Instruktionen haben. Sie würden dabei drei Gruppen von Probanden jeweils eine der Instruktionen geben und dann den Versuch zum Konstanzverfahren durchführen lassen. Das Ergebnis wäre, dass das Unsicherheitsintervall bei Instruktion 1 sehr klein und bei Instruktion 2 sehr breit ist. Die Breite des Unsicherheitsintervalls bei Instruktion 3 läge in der Mitte.

Vergleichen Sie jetzt Ihr Unsicherheitsintervall mit dem anderer Studenten! Sie werden feststellen, dass die Unterschiede recht groß sind. Dies kann einerseits die unterschiedliche Empfindlichkeit für die Wahrnehmung der Tonhöhe widerspiegeln. Es kann aber auch daran liegen, dass Sie selbst ein Typ sind, der eher nach Instruktion 1, Instruktion 2 oder Instruktion 3 urteilt. Man könnte mit diesem Versuch also auch so etwas wie „Urteilerpersönlichkeit“ erfassen. Wir haben versucht, eine eher neutrale Instruktion (entsprechend Instruktion 3) zu geben. An der Unterschiedlichkeit der Ergebnisse sieht man aber, dass dies nur teilweise gelungen ist. Man sieht an diesem Beispiel erneut, wie wichtig eine genaue Instruktion ist, um einheitliche Ergebnisse zu erhalten. Die Unterschiede zwischen den Probanden, die wir gefunden haben, sind nicht nur unterschiedliche Empfindlichkeiten in der akustischen Wahrnehmung, sondern können auch zum Teil auf unterschiedliche Arten zu urteilen zurückgeführt werden.

4.6. Zusammenfassung

Am Anfang dieses Kapitels hatten wir uns die Frage gestellt, welche Dimensionen oder Qualitäten eigentlich die Metrik des Psychischen ausmachen. Diese Frage konnten wir nicht beantworten, sondern haben festgestellt, dass diese Frage ein ganz wichtiges Thema der Psychologie ausmacht: Welche Arten des Erlebens gibt es überhaupt? Und: Gelten für diese verschiedenen Erlebensbereiche die gleichen Gesetzmäßigkeiten?

Um hier einen ersten Hinweis zu bekommen, hatten wir den Sinnesbereich gewechselt vom Sehen zum Hören, von der visuellen zur akustischen Wahrnehmung. Gleichzeitig haben wir eine neue Methode kennengelernt, um die Frage der Messgenauigkeit zu beantworten. Beim Konstanzverfahren werden verschiedene Vergleichsreize jeweils mit einem Standardreiz verglichen, wobei diese Vergleichsreize dem Standardreiz unterschiedlich ähnlich sind. Auf diese Weise lässt sich die Unterschiedsschwelle sehr genau bestimmen. Sie haben aber auch selbst erlebt, dass dieses Verfahren sehr aufwändig und für den Probanden recht unangenehm ist. Inhaltlich zeigt sich, dass die Unterschiedsschwelle für Tonhöhen wesentlich kleiner ist als die für Streckenlängen. Menschen können Töne unterschiedlicher Frequenz sehr gut unterscheiden.

Ganz am Ende hatten wir aber auch gesehen, dass die Ergebnisse von der Einstellung der Probanden abhängen, was man über die Instruktion teilweise beeinflussen kann. Dies zeigt, dass die Methode, mit der man vorgeht, die Ergebnisse beeinflussen kann. Die Diskussion über unterschiedliche Definitionen des Punkts subjektiver Gleichheit zeigten ebenfalls, dass unterschiedliche Methoden zu unterschiedlichen Messungen führen können. Die Frage nach der Metrik des Psychischen erfordert es daher nicht nur, diese in unterschiedlichen Erlebensbereichen zu bestimmen, sondern sich auch über Methoden Gedanken zu machen, mit denen man diese Ergebnisse erreichen kann. Diese Überlegung soll im folgenden Kapitel vertieft werden.

5. Versuch Grenzverfahren: Methode und Erkenntnis

5.1. Metrik und Methode

Wenn Sie den Titel dieses Kapitels lesen – was stellen Sie sich vor? Um was soll es gehen? Wir hatten das Thema schon ganz kurz erwähnt, als es um die Bedeutung der Instruktion ging. Als Psychologe sollte man wissen, wie sehr das eigene Wissen, die „Erkenntnis“, von der Methode abhängt. Ganz am Anfang des Buches ging es kurz darum, woher man als Wissenschaftler eigentlich etwas weiß. Die Psychologie, die Sie an der Universität kennenlernen, ist eine empirische Wissenschaft. Durch Beobachtung, durch Sammlung von Daten erkennen wir, wie das Erleben und Verhalten des Menschen funktioniert. Wir hatten dann an mehreren Beispielen im Selbstversuch analysiert, wie ganz einfache Aspekte des Erlebens funktionieren: Das Schätzen von Münzgrößen aus der Erinnerung, das Erleben der Streckenlänge und der Tonhöhe. Dabei waren unterschiedliche Methoden eingeführt worden: Das aktive Herstellen der gewünschten Größe und die Teilung von Strecken in zwei Hälften waren Herstellungsverfahren. Der Vergleich von Vergleichsreizen mit einem Standardreiz war das Konstanzverfahren. Mit diesen beiden Methoden konnten wir jeweils die Frage beantworten, wie genau unsere Wahrnehmung arbeitet. Durch die Berechnung der Unterschiedsschwelle wird dies beschrieben. Dadurch haben wir eine wichtige Gesetzmäßigkeit der Allgemeinen Psychologie kennengelernt: Das Weber'sche Gesetz. Es beschreibt, dass die Unterscheidungsfähigkeit für Reize von der Reizgröße abhängt. Beim Weber-Bruch wird die Unterschiedsschwelle durch die Reizgröße geteilt und beschreibt, wie empfindlich das untersuchte Sinnessystem ist. Wir hatten diesen methodischen Aspekt als „Metrik der Psychologie“ bezeichnet:

- Metrik der Psychologie: Welche Dimensionen (Erlebensbereiche) werden auf welche Weise (Messgenauigkeit, Art der Abbildung) erlebt?

Fast nebenbei war aber auch aufgefallen, dass die Ergebnisse teilweise schwierig zu interpretieren sind. Im ersten Versuch hatte sich gezeigt, dass kleine Münzen in der Größe unterschätzt werden, während größere realistischer eingeschätzt werden. Wir mussten ein zweites (kleines) Experiment durchführen, das darauf hinwies, dass hier nicht die Größe entscheidend ist, sondern der Wert von Münzen. Kleine Kreise wurden jedenfalls realistisch eingeschätzt. Dadurch folgerten wir, dass nicht die Größe, sondern der Wert eines Objekts die Größenschätzung aus dem Gedächtnis wesentlich beeinflusst. Aber kann man tatsächlich aus diesem einfachen Experiment eine solche weitreichende Folgerung ziehen? Werden Menschen, die man hoch achtet (z.B. Mahatma Gandhi) größer geschätzt als Menschen, die man wenig achtet (wer wäre hier ein gutes Beispiel)? Gilt dies auch für andere Eigenschaften wie das Gewicht? Wird das Gewicht von Goldbarren überschätzt, das von Steinbrocken unterschätzt?

Beim Versuch zum Konstanzverfahren hatte sich gezeigt, dass die Unterschiedsschwelle von der Instruktion abhängt und möglicherweise von der Persönlichkeit. Bei dem von uns durchgeführten Versuch war schwer zu sagen, ob eine große Unterschiedsschwelle durch eine schlechte Unterscheidungsfähigkeit für Tonhöhen entstanden war oder eine bestimmte Einstellung zum Versuch.

Beide Interpretationsprobleme entstehen dadurch, dass verschiedene mögliche Einflussgrößen gleichzeitig wirken. Wie kann man die Datensammlung so gestalten, dass die Interpretation möglichst eindeutig wird? Hier geht es um einen weiteren Aspekt der Methodik, bei dem „Versuchsplanung“ und „Experiment“ wichtige Stichpunkte sind. Nur mit Hilfe einer sauberen Methodik kann es gelingen, wissenschaftliche, belastbare Ergebnisse zu erhalten. Darum soll es in diesem Kapitel gehen.

- Methode der Psychologie: Auf welche Weise muss eine wissenschaftliche Untersuchung gestaltet werden, um aussagekräftige Erkenntnisse zu erhalten?

5.2. Die experimentelle Methode

Die Grundidee des Experiments war bereits in Kapitel 2.5.2 auf Seite 20 dargestellt worden. Als Psychologe sind wir an Gesetzmäßigkeiten der Form

- $AV = f(UV)$

oder

- $UV \rightarrow AV$

interessiert. Dahinter steckt das S-O-R-Schema: Bestimmte Reize (S) wirken auf den Organismus (O), der darauf mit einem bestimmten Verhalten oder Erleben (R) reagiert. Man kann damit erklären, warum ein bestimmtes Verhalten oder Erleben gezeigt wird: Jemand hat Angst vor großen Höhen, weil er einmal ein traumatisches Erlebnis in großer Höhe erlebt hat. Man kann damit umgekehrt auch versuchen, das Verhalten zu ändern: Wenn ich einen Patienten mit Höhenangst immer wieder großen Höhen aussetze (eine sogenannte Expositionstherapie), wird er erleben, dass nichts Negatives passiert und so seine Ängste wieder verlieren. Die abhängige Variable (AV) „Höhenangst“ ist also eine Funktion der unabhängigen Variablen (UV) „Angsterleben in großer Höhe“ ($AV = f(UV)$). Oder: Man möchte beweisen, dass durch ein Angsterleben in großer Höhe Höhenangst entstehen kann ($UV \rightarrow AV$). Aber wie beweist man dies wissenschaftlich?

1) Planmäßigkeit: Beim Experiment stellt man aktiv und gezielt eine Ursache-Wirkungs-Beziehung her. Man wartet nicht, bis jemand aus Zufall in großer Höhe Angst bekommt und untersucht, ob er eine Höhenangst entwickelt. Man versucht vielmehr gezielt, einer Reihe von Personen in großer Höhe Angst einzuflößen. Man könnte sie z.B. auf einen hohen Balkon bringen und sie dann hinter ihrem Rücken mit einem lauten Knall erschrecken. Eine Woche später könnte man sie erneut auf den Balkon einladen und dann messen, inwieweit sie Höhenangst erleben (dieses Beispiel klingt vielleicht seltsam, aber John B. Watson wurde mit seinem ganz ähnlich aufgebauten Versuch mit dem kleinen Albert damit berühmt).

Gelingt dieser Nachweis, spricht bereits einiges dafür, dass der Schreck zu Höhenangst führen kann. Allerdings muss man sicher sein, dass wirklich nur der Schreck dafür verantwortlich war. Möglicherweise haben besonders schreckhafte Probanden an dem Versuch teilgenommen. Vielleicht war der Balkon baufällig, so dass dadurch Angst entstand. Man muss also auf eine Weise vorgehen, die alternative Interpretationen ausschließt.

2) Isolierende Variation: Das Experiment wird so gestaltet, dass nur der relevante Einflussfaktor wirken kann und alle anderen möglichen Einflussfaktoren kontrolliert werden. Um sicherzugehen, dass nur der Schreck zur Höhenangst führt, würde man eine zweite Gruppe von Probanden ebenfalls auf den Balkon bringen, dort aber nicht erschrecken. Ansonsten würde man sie möglichst gleich wie die erste Gruppe behandeln. Um den Einfluss unterschiedlicher Persönlichkeiten auszuschließen, würde man die Probanden per Zufall entweder der einen oder anderen Gruppe zuweisen. Auf diese Weise sollten die beiden Gruppen vor Beginn des Versuchs vergleichbar sein. Wenn dann eine Woche später die Gruppe mit dem Knall Anzeichen für Höhenangst zeigt, die andere Gruppe nicht, kann man durchaus schon recht sicher sein, dass der Knall tatsächlich die Ursache für die Höhenangst ist.

Diese ersten beiden Überlegungen führen zum Grundschema des Experiments (s. Abbildung 24): Die Versuchspersonen werden per Zufall in eine von zwei Gruppen eingeteilt. Man nennt die „Randomisierung“. Durch diese zufällige Zuteilung wird erreicht, dass die beiden Gruppen vor Beginn des Versuchs und bevor die unabhängige Variable wirken kann, vergleichbar sind. Die Gruppe, in der die Einflussgröße wirken soll, wird als „Experimentalgruppe“ bezeichnet. Im Beispiel ist dies die Gruppe von Versuchspersonen, die auf dem Balkon erschreckt wird.

Diese wird mit einer zweiten Gruppe verglichen, die genauso behandelt wird, also auch auf den Balkon gebracht wird. Diese Versuchspersonen werden aber nicht erschreckt. Man bezeichnet diese Gruppe als „Kontrollgruppe“. Bei beiden Gruppen misst man dann die Wirkung, die durch die UV entstehen sollte. Im Beispiel ist dies die Höhenangst. Ist nur in der Kontrollgruppe die Höhenangst geringer als in der Experimentalgruppe, so muss dafür der Knall verantwortlich gewesen sein. Denn dies ist der einzige Unterschied zwischen den Gruppen und nur so ist zu erklären, dass die Experimentalgruppe eine stärkere Höhenangst hat. Der experimentelle Beweis enthält eigentlich zwei Nachweise: In der Experimentalgruppe entsteht Höhenangst, in der Kontrollgruppe nicht. Dies kann man auch noch überzeugender nachweisen, wenn man auch vor dem Versuch, also bevor der Knall verabreicht wird, zunächst die Höhenangst misst. Durch die zufällige Zuweisung kann man zwar davon ausgehen, dass die Gruppen vergleichbar sein müssten, aber durch eine solche Vorher-Messung beweist man das noch. Man kann dann auch noch in jeder Gruppe untersuchen, ob sich die Höhenangst wirklich von vorher zu nachher verändert hat.

	UV	AV
Zufall	Experimentalgruppe: Mit Knall	Höhenangst
	Kontrollgruppe: Ohne Knall	

Abbildung 24: Grundschemata des Experiments für das Beispiel im Text.

Ganz auszuschließen ist natürlich nicht, dass ein solches Ergebnis an wenigen Probanden auch einmal zufällig entsteht. Vielleicht sind trotz der zufälligen Zuteilung doch mehr ängstliche Probanden in der einen Gruppe. Vielleicht war zufällig in der einen Gruppe das Wetter schlecht, so dass sie Angst bekamen, auszurutschen und vom Balkon zu fallen. Man muss also sichergehen, dass das Ergebnis des Experiments nicht ein zufälliges, singuläres Ereignis ist.

3) Wiederholbarkeit: Experimente müssen wiederholt werden. Am besten ist es, wenn eine andere Arbeitsgruppe die Experimente mit dem gleichen Ergebnis wiederholen kann. Man nennt dies eine Replikation. Je öfter ein Experiment repliziert werden kann, umso stärker spricht dies für die Gesetzmäßigkeit, die hier gezeigt werden soll. Dieser Grundsatz der Wiederholbarkeit ist auch der Grund dafür, dass in Berichten über Experimente (wissenschaftliche Zeitschriftenartikel) die Methodik so ausführlich dargestellt wird. Andere Wissenschaftler müssen in der Lage sein, das Experiment zu wiederholen (zu replizieren).

Damit sind die Grundprinzipien des Experiments ganz kurz dargestellt worden. Wir werden dies in den folgenden Kapiteln noch vertiefen. Wichtig ist, dass wir mit diesem experimentellen Vorgehen prüfen können, inwieweit bestimmte Einflussgrößen systematisch das Erleben und Verhalten beeinflussen. Eine solche Einflussgröße kann auch die Messmethodik sein: Mit einem Experiment kann man systematisch untersuchen, ob eine Größenschätzung mit zwei verschiedenen Messmethoden zu unterschiedlichen Ergebnissen führt. Darum geht es im folgenden Abschnitt.

5.3. Die Messmethode als unabhängige Variable

Sie hatten in den ersten beiden Versuchen das Herstellungsverfahren kennengelernt, bei dem Sie selbst aktiv eine bestimmte Größe (Münzen) oder Länge (Streckenteilung) hergestellt haben. Im dritten Versuch hatten Sie das Erleben der Tonhöhe mit dem Konstanzverfahren untersucht. Dabei zeigte sich, dass die Unterscheidungsfähigkeit für Tonhöhen sehr viel höher ist als die für die Länge von Strecken. Allerdings hatten wir nicht nur die beiden verschiedenen Sinnessysteme untersucht, sondern gleichzeitig die Messmethode gewechselt. Damit wurde gegen

das Prinzip der isolierenden Variation verstoßen: Als UV wurde gleichzeitig Messmethode und Sinnessystem variiert. Beide Einflüsse sind deshalb nicht zu trennen. Damit wäre es prinzipiell auch möglich, dass die höhere Empfindlichkeit für Tonhöhe gar nicht durch das Sinnessystem bestimmt ist, sondern, dass Sie mit dem Konstanzverfahren bereits sehr kleine Unterschiede besser wahrnehmen können als mit dem Herstellungsverfahren. Zwischen diesen beiden Alternativen kann man wissenschaftlich nur unterscheiden, indem man diese Einflussgrößen getrennt untersucht.

Im folgenden Versuch soll der Einfluss der Methode bestimmt werden. Ändert sich die Unterschiedsschwelle, wenn diese mit unterschiedlichen Methoden gemessen wird? Die Unterschiedsschwelle hatten wir beim Streckenmitteln für verschiedene Streckenlängen gemessen, indem man diese halbiert hatte und sich dabei in mehreren Schritten auf die richtige Länge eingependelt hatte. Im folgenden Versuch wird eine etwas andere Methode eingeführt, das Grenzverfahren. Das Vorgehen ist in Abbildung 25 für das Streckenmitteln dargestellt.

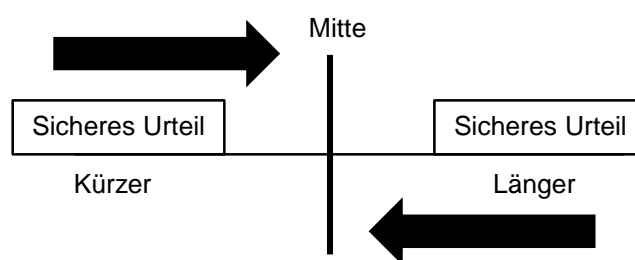


Abbildung 25: Vorgehen beim Grenzverfahren. Beim Streckenmitteln nähert man sich entweder von links (deutlich kürzer) oder von rechts (deutlich länger) der Mitte an.

Wie in der Abbildung dargestellt, beginnt man z.B. auf der linken Seite, bei der die linke Strecke deutlich kürzer ist. Man nähert sich dann langsam der Mitte an und hält an, wenn man diese erreicht hat. Man könnte das auch von rechts machen, um sicherzustellen, dass die Seite des Herangehens keine Rolle spielt. Der wesentliche Unterschied zum Herstellungsverfahren liegt also darin, dass man sich nicht in mehreren Schritten auf die Mitte einpendeln kann. Dafür geht der Versuch deutlich schneller, da man für jedes Urteil nur eine Annäherung benötigt. Die Frage unseres Versuchs ist nun, ob diese etwas andere Messmethode die Ergebnisse verändert. Um diese Frage zu beantworten, lesen Sie bitte die folgende Instruktion und führen Sie den Versuch durch.

5.4. Instruktion

In dem Versuch geht es um das Einschätzen von Streckenlängen. Sie werden wieder die 7 verschiedenen lange Strecken erhalten, die sie jeweils halbieren sollen. Dazu erhalten Sie einen Cursorbalken, den Sie auf der Strecke bewegen können. Sie beginnen immer ganz links und bewegen dann den vertikalen Balken in die Mitte. Nähern Sie sich jetzt langsam und vorsichtig der Mitte und halten Sie an und drücken die linke Maustaste, wenn Sie diese erreicht haben. Dann ist der Durchgang beendet und die nächste Strecke erscheint. Jede der sieben Streckenlängen wird fünf Mal beurteilt.

Wichtig ist, dass Sie dabei langsam und sorgfältig vorgehen, denn Sie können Ihr Urteil nicht mehr verändern. Achten Sie also darauf, dass Sie rechtzeitig anhalten und nicht über die Mitte hinauschießen! Wenn Sie merken, dass Sie viel zu weit gegangen sind und die Messung damit fehlerhaft ist, merken Sie sich das und Sie können das dann in der Excel-Tabelle korrigieren.

Wenn Sie alles verstanden haben, starten Sie bitte das Programm und den Versuch „Grenzverfahren“. Sie sehen auf der Instruktionsseite noch einmal diese Informationen und können dann

beginnen. Am Ende erscheint eine Tabelle mit Ihren Ergebnissen, die Sie nach Excel kopieren können.

Wie bei jedem dieser Versuche ist es wichtig, dass Sie konzentriert und ungestört arbeiten. Denken Sie nicht lange über Ihre Urteile nach, sondern mitteln Sie die Strecken so sorgfältig wie möglich. Bitte starten Sie nun den Versuch!

Und hier geht es zum Versuch:

- <https://cgi-psych.rz.tu-bs.de>

5.5. Auswertung und Ergebnis

Die Versuchsfrage war, ob die Messmethode die Ergebnisse verändert. Dies bezog sich vor allem auf die Messgenauigkeit, die wir mit Hilfe der Unterschiedsschwelle bzw. des Unsicherheitsintervalls berechnet hatten. In Kapitel 3.4.3 auf Seite 30 hatten wir für das Herstellungsverfahren für jede Streckenlänge Mittelwert und Standardabweichung berechnet. Die Standardabweichung gibt an, wie sehr die Urteile bei einer Streckenlänge streuten. Sie wurde deshalb als Maß für die Unterschiedsschwelle verwendet. Der Mittelwert \pm die Standardabweichung ergab das Unsicherheitsintervall. Sie finden zur Auswertung der Daten des Grenzverfahrens genau dieselben Berechnungen wie beim Herstellungsverfahren (Versuch Streckenmitteln).

5.5.1. Die Messgenauigkeit

Die zentrale Frage des Versuchs war, wie gut die Längenschätzung mit dem Grenzverfahren funktioniert. Wie wir oben beim Streckenmitteln mit dem Herstellungsverfahren gesehen hatten, ist die Unterschiedsschwelle US ein Maß für die Genauigkeit der Mittelung. Gelingen die Streckenmittelung und damit die Längenschätzung mit dem Grenzverfahren mit derselben Genauigkeit wie mit dem Herstellungsverfahren? Um dies zu berechnen, berechnen wir für die Durchgänge die Standardabweichung. Die Ergebnisse für 18 studentische Probanden sind in Abbildung 26 dargestellt. Man erkennt auch bei dieser Methode sehr schön das Weber'sche Gesetz: Mit zunehmender Reizgröße nimmt die Unterschiedsschwelle zu.

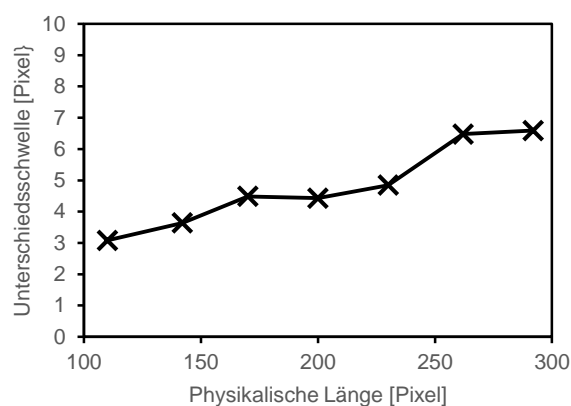


Abbildung 26: Unterschiedsschwelle (Mittelwert der individuellen Standardabweichungen der 18 studentischen Probanden) in Abhängigkeit von der Streckenlänge.

Berechnen Sie dies nun auch für Ihre eigenen Daten. In der Excel-Tabelle finden Sie dazu die Vorlage, in der Sie die Berechnungen vornehmen können. Sie können hier auch wieder eine Gerade anpassen.

Berechnet man jetzt den Weber-Bruch (Unterschiedsschwelle geteilt durch Streckenlänge), so ergibt sich annähernd eine Konstante, wie Abbildung 27 zeigt. Bei der Stichprobe der studentischen Probanden liegt diese bei 0.025. Berechnen Sie für Ihre eigenen Daten ebenfalls den Weber-Bruch. Dazu müssen Sie für jede Streckenlänge den Mittelwert der beiden Standardabweichungen (als Maß für die Unterschiedsschwelle) durch die gemessene Streckenlänge teilen. Sie kennen die entsprechenden Zeilen und Berechnungen ja schon vom Streckenmitteln mit dem Grenzverfahren und finden dies in der Excel-Tabelle auch wieder.

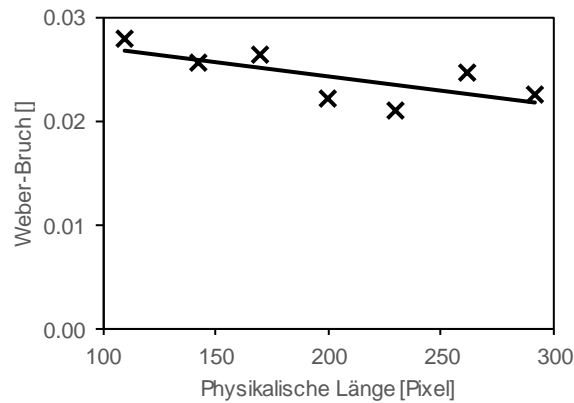


Abbildung 27: Der Weber-Bruch in Abhängigkeit von der physikalischen Länge. Dargestellt ist die Unterschiedsschwelle geteilt durch die Strichlänge.

Jetzt kommt der spannende Augenblick: Ist dieser Wert, gemessen mit dem Grenzverfahren, anders als bei der Messung mit dem Herstellungsverfahren? Schauen Sie in Abbildung 16 auf Seite 35 nach und vergleichen Sie diese mit Abbildung 27! Und vergleichen Sie Ihre eigenen Daten.

Tatsächlich ist der Weber-Bruch beim Grenzverfahren etwas größer als beim Herstellungsverfahren. Allerdings fällt bei genauer Betrachtung der Daten auf, dass die Abweichungen bei einem Probanden entweder eher negativ sind oder eher positiv. Dies ist eine Frage der Strategie:

- Wenn man von links kommt, klickt man sobald man den Eindruck hat, dass man an der Mitte angekommen ist, lieber etwas früher als zu spät. Dann erhält man insgesamt negative Abweichungen.
- Wenn man von links kommt, klickt man erst, wenn man ganz sicher ist, dass man auf der Mitte ist, und lieber etwas später als zu früh. Dann erhält man insgesamt positive Abweichungen.

Sehen Sie mal in Ihre Daten, und versuchen Sie sich zu erinnern, welche Strategie Sie verfolgt haben. Es ist typisch für das Grenzverfahren, dass man die Mitte entweder etwas überschätzt oder unterschätzt. Wenn man diese Strategie bei allen Strecken anwendet, ist die Genauigkeit ähnlich hoch wie beim Herstellungsverfahren. Das Ergebnis ist aber insofern etwas ungenauer, dass eben nicht genau die Mitte getroffen wird, sondern man entweder etwas links oder rechts von der Mitte landet. Man führt deshalb gerne das Grenzverfahren einmal von links und einmal von rechts durch, und mittelt dann die Werte.

Ist Ihnen beim letzten Absatz methodisch eigentlich etwas aufgefallen? Man hätte den Versuch verbessern können, indem man die Probanden instruiert, eine der beiden Strategien anzuwenden. Dann wären die Unterschiede zwischen den Probanden nicht so groß und man wüsste, dass entweder eine leichte Unterschätzung oder Überschätzung herauskommt. Man könnte auch diese Strategien erklären und die Probanden bitte, weder die eine noch die andere Strategie zu verwenden. Ist das möglich und führt dies zu besseren Ergebnissen? Man könnte das empirisch

prüfen, indem man die Instruktion variiert. Vielleicht haben Sie ja auch Lust, das mit sich selbst zu probieren!

5.5.2. Zusammenfassung der Ergebnisse

Fassen wir also die Ergebnisse dieses Versuchs zusammen:

- Je nach Messmethode kann die Messgenauigkeit des Erlebens unterschiedlich erscheinen: Mit dem Herstellungsverfahren erscheint die Messgenauigkeit höher als mit dem Grenzverfahren, da eine Unterschätzung oder Überschätzung vermieden wird. Auf der anderen Seite geht das Grenzverfahren etwas schneller, da man sich nicht einpendeln muss. Wiederum andererseits könnte das die Probanden demotivieren, da man sein Urteil nicht mehr korrigieren kann.
- Zusammenfassend ist das Herstellungsverfahren für den Probanden angenehmer und liefert genauere Ergebnisse. Das Grenzverfahren geht dafür etwas schneller.

Warum einigt man sich dann nicht auf eine Methode? Warum sind verschiedene Methoden notwendig? Wenn es um die grundlegende Empfindlichkeit der Wahrnehmung in verschiedenen Erlebensbereichen geht, ist es tatsächlich sinnvoll, sich auf eine Methode zu einigen, um auf diese Weise die unterschiedliche Messgenauigkeit der Sinnessysteme abzubilden – unabhängig von den störenden Einflüssen unterschiedlicher Methoden. Wenn es allerdings um praktische Anwendungen geht, muss die Methodik der Fragestellung angepasst werden.

Wenn Sie zum Beispiel messen wollen, ab welcher Lichtstärke ein Verkehrsschild bei Nacht in einem bestimmten Abstand gesehen wird, um damit Lichtstärken für Scheinwerfer festzulegen, müssen Sie eine Absolutschwelle bestimmen. Sie könnten das mit dem Herstellungsverfahren machen und die Probanden in einer solchen Situation das Scheinwerferlicht so einstellen lassen, dass sie das Schild gerade erkennen können. Sie könnte auch das Konstanzverfahren anwenden und bei verschiedenen Lichtstärken jeweils beurteilen lassen, ob das Schild gesehen wird oder nicht. Beide Verfahren haben den Nachteil, dass Sie damit wahrscheinlich zu geringe Lichtstärken festlegen. Die Probanden wissen schon, an welche Stelle das Schild steht und werden dann schon mit „ja“ antworten, wenn sie irgendeine Veränderung wahrnehmen. Das würde im realen Verkehr aber noch lange nicht ausreichen, um tatsächlich ein Schild erkennen zu können. Hier erscheint das aufsteigende Grenzverfahren der realen Situation des Fahrers in der Nacht vergleichbarer. Der Proband wartet im Dunkeln, während der Versuchsleiter langsam die Lichtstärke ansteigen lässt. Sobald er dann an irgendeiner Stelle das Verkehrsschild erkennt, soll der Proband antworten. Das führt zu einer sehr ähnlichen Situation wie bei einer Nachtfahrt, so dass die mit dieser Methode gefundenen Werte sehr viel sinnvoller für diese Fragestellung sind.

- Je nach Fragestellung sind unterschiedliche Methoden „am besten“. Wenn Sie eine eigene Untersuchung planen oder eine Beschreibung einer Studie lesen, sollten Sie sich deshalb immer fragen: Ist die Methode, die gewählt wurde, tatsächlich am besten zur Beantwortung der Frage geeignet?

5.6. Zusammenfassung und Ausblick

5.6.1. Methoden und Phänomene

In diesem Kapitel geht es um den Zusammenhang zwischen Methode und Erkenntnis. Jede wissenschaftliche Erkenntnis ist abhängig von der Methode, mit der diese gewonnen wurde. Das zeigte sich auch bei dem ganz einfachen Wahrnehmungsexperiment, was hier durchgeführt wurde. In Kapitel 3 hatten wir die Messgenauigkeit für die Wahrnehmung von Streckenlängen

mit dem Herstellungsverfahren untersucht, jetzt mit dem Grenzverfahren. Führt die andere Methode zu anderen Ergebnissen? Bei der Diskussion der Ergebnisse hatten wir mehrere methodische Einflussfaktoren gefunden:

- Die Urteilsstrategie der Probanden, die auch durch Instruktion zu verändern ist, beeinflusst vor allem die Bestimmung des Punkts subjektiver Gleichheit.
- Die Messmethode (Herstellungsverfahren vs. Grenzverfahren) führt bei gleicher Berechnung der Unterschiedsschwelle (Standardabweichung über alle Durchgänge) zu einer unterschiedlichen Abschätzung der Messgenauigkeit: Mit dem Herstellungsverfahren erscheint die Unterschiedsschwelle etwas kleiner.

Zusammenfassend wurde so sehr deutlich, dass das Ergebnis von der Methode abhängt. Dies bezeichnet man auch als „Methodenabhängigkeit der Phänomene“. Daraus folgt, dass man die Methode kennen muss, um die Ergebnisse interpretieren zu können. Wurden wirklich dieselben Methoden verwendet, so dass man die Ergebnisse zweier Studien vergleichen kann? Was wurde überhaupt gemessen? Welche Aussage erlaubt das? Wie wurden die Probanden instruiert? Wie sind sie an die Aufgabe herangegangen? Erst wenn man diese Fragen beantworten kann, kann man die Ergebnisse wirklich verstehen. Deshalb ist es so wichtig, in wissenschaftlichen Darstellungen die Methode ausführlich zu beschreiben. Außerdem ist dies die Voraussetzung dafür, das Experiment replizieren zu können. Diese Wiederholbarkeit hatten wir bereits als ein wesentliches Merkmal der wissenschaftlichen Methodik kennengelernt.

Umgekehrt lässt sich daraus auch ableiten, dass je nach Untersuchungsgegenstand die richtige Methode gewählt werden muss. Dies bezeichnet man als „Phänomenabhängigkeit der Methode“. Am Beispiel der Beleuchtungsstärke, die man benötigt um ein Verkehrsschild nachts zu sehen, war gezeigt worden, dass in dieser Situation das Grenzverfahren angemessen ist, während Konstanz- und Herstellungsverfahren weniger geeignet schienen. Bei der Beurteilung wissenschaftlicher Ergebnisse, aber vor allem auch bei der Planung eigener Studien, ist die erste Frage damit: Wie kann ich das messen, was mich interessiert?

5.6.2. Ausblick: Variationen der Messmethoden

Wir haben bislang im Selbstversuch drei Messmethoden kennengelernt: Herstellungsverfahren, Konstanzverfahren und Grenzverfahren. Bisher haben wir nicht diskutiert, dass bei jedem dieser Verfahren unterschiedlich vorgegangen werden kann. Einige Beispiele für Variationen sind folgende:

- Die aktive Variation des Reizes durch den Probanden beim Herstellungsverfahren vs. die gezielte Variation durch den Versuchsleiter
- Die Wahl unterschiedlich feiner Abstufungen beim Herstellungsverfahren
- Die Anzahl der Vergleichsreize beim Konstanzverfahren
- Die Reihenfolge des Vergleichs beim Konstanzverfahren (Standardreiz, dann Vergleichsreiz vs. umgekehrt)
- Die Reihenfolge der Reize beim Konstanzverfahren
- Senkrechte vs. waagerechte Striche beim Streckenmitteln

Diese Liste ließe sich fast beliebig verlängern. Für eine ganze Reihe solcher Variation wurde bereits untersucht, in welcher Weise sie sich auf die Ergebnisse auswirken. Sie lassen sich in drei Gruppen einteilen, für die im Folgenden einige Beispiele genannt werden:

- Zeitliche Einflüsse: Die Abfolge von Standardreiz und Vergleichsreiz, die Dauer der Darbietung der Reize und die Pausendauer zwischen zwei Reizen, die verglichen werden sollen.
- Räumliche Einflüsse: Die Anordnung der Reize in horizontal und vertikal, der Abstand des Probanden vom Bildschirm und die Annäherung von rechts oder links beim Grenzverfahren.
- Serielle Einflüsse: Die Reihenfolge (z.B. als ersten Reiz einen sehr großen vs. einen sehr kleinen Reiz), die Anzahl der Reize, die beurteilt werden und die Verteilung der Vergleichsreize in Relation zum Standardreiz.

Diese Arten von Einflüssen sind durch den Versuchsleiter zu kontrollieren, indem die Messung auf eine bestimmte Weise gestaltet wird. Dabei ist natürlich die Phänomenabhängigkeit der Methode zu beachten: Ist dieses Vorgehen der möglichen praktischen Anwendung angemessen?

Hinzu kommen Einflüsse im Probanden, die zum Teil durch die Instruktion zu kontrollieren sind. Wie soll das Urteil gefällt werden? Was bedeutet „gleich“? Wie soll man bei dem Versuch vorgehen? Außerdem sind Veränderungen zu beachten, die nur bedingt zu kontrollieren sind: Die Probanden werden im Laufe der Zeit müder (Versuchsdauer anpassen) oder das Interesse am Versuch ist gering oder sinkt, weil der Versuch langweilig ist (durch Geld motivieren?). Vielleicht handelt es sich aber auch um einen vorsichtigen Menschen, der sich ungern festlegt. Das werden Sie auch durch eine Instruktion nur wenig ändern können.

Auch diese Einflüsse zeigen, wie sehr die Methode die Ergebnisse beeinflussen kann. Es ist daher wichtig, im Laufe des Studiums eine „Methodenkompetenz“ in diesem Sinne zu erwerben. Man sollte wissen, welche Auswirkungen es hat, wenn man bestimmte Details im Vorgehen verändert. Auf diese Weise kann man gezielt die für den jeweiligen Untersuchungsgegenstand „beste“ Methode auswählen.

5.6.3. Ein kleiner philosophischer Ausblick

Alles ist relativ. Dieser Eindruck könnte nach dem Lesen dieses Kapitels entstehen. Wenn man die Methodenabhängigkeit der Phänomene ins Extrem treibt, könnte das bedeuten, dass man „das“ Erleben des Menschen gar nicht untersuchen kann, weil ja je nach Methode die Ergebnisse ganz unterschiedlich sind. Wenn man ein strenger Wissenschaftler in diesem Sinne ist, wird man nicht mehr von einem bestimmten Erleben sprechen („Ich habe Intelligenz untersucht!“), sondern nur die Methode beschreiben („Ich habe untersucht, was mein Intelligenztest misst!“). Diese Richtung bezeichnet man als Operationalismus: Man möchte zwar bestimmte Phänomene wie Intelligenz untersuchen, ist aber überzeugt davon, dass man das Phänomen selbst sowieso nicht untersuchen kann. Aber man kann eine Methode angeben, die zumindest bestimmten Aspekten von Intelligenz nahekommt, z.B. einen Intelligenztest mit ganz bestimmten Aufgaben. Man hat damit „operationalisiert“, wie man dieses messen möchte, will aber auch deutlich machen, dass einem klar ist, dass es sich hier nicht um „die“ Intelligenz handelt. Also: Man untersucht nicht Intelligenz, sondern das, was der Intelligenztest misst.

Richtig ist an diesem Ansatz, dass er die Methodenabhängigkeit der Phänomene ernst nimmt. Eine Aussage über psychische Phänomene ist nur so gut, wie die Methode, mit der man dieses Phänomen gemessen hat. Allerdings scheint mir die Sicht, dass man so nie das Phänomen messen kann, zu pessimistisch. Am ganz einfachen Beispiel unseres Versuchs: Natürlich hängen die Aussagen über die erlebte Mitte und die Messgenauigkeit dieses Erlebens von der Messmethode ab. Deshalb zu leugnen, dass es so etwas wie ein Erleben der Mitte gibt, erscheint absurd. Wenn ich die Messmethode konstant halte, erscheint auch die Aussage, dass die Messgenauigkeit beim Erleben von Streckenlängen kleiner ist als beim Erleben von Tonhöhen, sinnvoll. Der

Zugang zu den psychischen Phänomenen ist aus dieser Sicht bestimmt durch die Methodik, macht diesen Zugang aber nicht unmöglich. Und außerdem: Diese Begrenzung gilt für alle Wissenschaften, insbesondere für die Physik. Ist Licht eine Welle oder ein Teilchen? Das hängt ab von der Messmethode. Die Psychologie unterscheidet sich hier also nicht vor anderen Naturwissenschaften.

6. Einschub: Kritische Bewertung von Artikeln

Das Seminar hat drei wesentliche Ziele:

- Sie sollen verstehen, was „Messen“ in der Psychologie bedeutet. Das Messen von Erleben ist die Voraussetzung, dass man als Psychologe Wissenschaft betreiben kann.
- Dabei werden Sie wesentliche Eigenschaften des Erlebens kennenlernen, ein wichtiger Teilbereich der Allgemeinen Psychologie.
- Sie lernen, wie man psychologische Gesetzmäßigkeiten nachweist. Das ist die experimentelle Methodik der Psychologie.

Experimente sind der Königsweg, um kausale Beziehungen nachzuweisen. In den bisherigen Versuchen haben Sie dazu schon einiges gelernt. Ein wesentlicher Teil der Psychologie als Wissenschaft ist es, neues Wissen auf Basis von experimentellen Studien zu erarbeiten und dem Rest der wissenschaftlichen Gemeinschaft zu präsentieren. Dies geschieht zunächst in Originalartikeln, in denen eine Fragestellung hergeleitet wird, eine Methode zu Beantwortung der Frage dargestellt wird und dann die Ergebnisse präsentiert und diskutiert werden.

Ein zentrales Lernziel im Studium an einer Universität ist die Fähigkeit, mit Wissen kritisch umzugehen. Sie lernen hier durch Selbstversuche, wie Fragen untersucht und beantwortet werden können. Wir möchten dies in dem Einschub noch ergänzen, indem Sie sich kritisch mit präsentiertem Wissen (in Form von Originalarbeiten) auseinandersetzen.

Dazu haben wir ein Beispiel vorbereitet. Wir werden dort jeweils sowohl inhaltlich als auch methodisch eine wissenschaftliche Arbeit analysieren und bewerten. Das folgende Kapitel gibt dazu einige Hinweise, wie man den Inhalt einer Arbeit am besten verstehen kann.

6.1. Inhalt lesen und verstehen

Eine psychologische Arbeit ist meist folgendermaßen aufgebaut:

- Abstract – Zusammenfassung
- Einführung – Fragestellung
- Methodik
- Ergebnisse
- Diskussion

In sehr kurzer Form wird zunächst der Inhalt des Artikels zusammengefasst. Der erste Schritt zum Verständnis ist damit das Lesen und Verstehen des Abstracts. Von der Lernforschung her ist es sinnvoll, zunächst einen groben Rahmen, eine Idee zu entwickeln, worum es in dem Artikel geht. Dies erleichtert und fördert das Verstehen der Details, wenn man den Artikel dann liest. Ein gutes Abstract macht verständlich, welches Thema untersucht wird, welche Bedeutung dieses Thema hat und was die Fragestellung der Arbeit ist. Man sollte außerdem ungefähr verstehen, wie die Untersuchung durchgeführt wurde (Methodik) und was das wesentliche Ergebnis war. Eine Bewertung der Ergebnisse und ein Ausblick auf ihre Relevanz schließt das Abstract ab.

In der Einführung wird zunächst erklärt, mit welchem Thema sich die Arbeit beschäftigt. Dabei wird auch erklärt, warum dieses Thema überhaupt wissenschaftlich relevant ist und welche Bedeutung es hat. Ganz wichtig ist dann die Erklärung, wo Unklarheiten oder offene Fragen liegen. Dazu wird ein kurzer Überblick über den Stand der Forschung gegeben, um demgegenüber zu zeigen, wo die eigene Arbeit ansetzt und offene Fragen klärt.

Im Bereich der Methodik wird dargestellt, wie diese Fragen untersucht wurden. Diese Erklärung sollte so gut sein, dass man die Untersuchung selbst durchführen könnte. Wie Sie in den Selbstversuchen gesehen hatten, ist es ganz entscheidend zu verstehen, was eigentlich die Aufgabe der Probanden war und wie ihr Erleben und Verhalten gemessen wurde. Dies muss in der Methodik entsprechend beschrieben werden.

Der Ergebnisteil stellt die Beantwortung der Fragestellung dar. In der Regel erfolgt eine statistische Prüfung, ob gefundene Unterschiede tatsächlich auch bedeutsam sind oder auch durch Zufall entstanden sein könnten. Wichtiger ist die Beschreibung, wie sich das Verhalten und Erleben geändert haben. Grafische Darstellungen der wichtigsten Befunde sind dabei hilfreich.

Schließlich werden die Ergebnisse in der Diskussion bewertet. Beantworten sie die eingangs gestellte Frage? Passen die Ergebnisse zu den bisherigen Befunden? Wie sind die Ergebnisse zu interpretieren? Versteht man damit den Menschen besser? Gibt es praktische Konsequenzen?

Nach diesem kurzen theoretischen Überblick soll es jetzt praktisch werden. Bitte lesen Sie jetzt den Artikel und versuchen Sie ihn zu verstehen. Beantworten Sie dazu in Stichpunkten die folgenden Fragen:

- Abstract: Um welches Thema geht es? Warum ist das Thema wichtig? Was war die Frage? Wie wurde sie untersucht? Was kam raus? Was bedeutet das?
- Einleitung: Warum ist das Thema relevant, warum beschäftigt man sich damit? Was sind wesentliche Erkenntnisse, die schon vorliegen? Was ist die offene Frage, die hier untersucht werden soll?
- Methodik: Wie wurde die Untersuchung durchgeführt?

An dieser Stelle eine ganz kurze Pause: Versuchen Sie erst mal, die Methode grob zu verstehen. Es soll im ersten Schritt darum gehen, den Artikel insgesamt zu verstehen. Im zweiten Schritt werden wir dann die Methodik im Detail anschauen. Diese Frage können Sie also ruhig zunächst sehr kurz beantworten.

- Ergebnisse: Was sind die wesentlichen Ergebnisse?
- Diskussion: Was bedeuten die Ergebnisse? Beantworten sie die Fragestellung des Artikels? Hat das weitere Konsequenzen?

Lesen Sie bis zur nächsten Stunde den Artikel und beantworten Sie diese Fragen. Wir werden sie dann gemeinsam diskutieren.

6.2. Methodik beschreiben und kritisieren

Bis zur nächsten Stunde sollen Sie nun die Methodik des Artikels etwas näher unter die Lupe nehmen. Beginnen Sie damit, indem Sie zunächst nach folgenden Informationen suchen:

- Versuchsplan: Was waren die UV? Welche Stufen wurden gewählt? Wie wurden diese genau umgesetzt? Warum? Wie wurden Störvariablen kontrolliert? Was ist wirklich der Unterschied zwischen den Stufen der UV? Welche Gruppen von Probanden wurden untersucht (Messwiederholung, unabhängige Gruppen)?
- Abhängige Variablen: Was wurde gemessen? Wie? Sind diese Maße sinnvoll? Wurde wirklich das gemessen, was psychologisch relevant ist?
- Instruktion: Wie wurden die Probanden instruiert? Was war ihre Aufgabe? Wie sollten sie an den Versuch herangehen?

- Durchführung und Ablauf: Wie war der Versuch aufgebaut, wie muss man sich den Ablauf vorstellen?
- Stichprobe: Welche Personen wurden untersucht? Wie wurden diese gewonnen? Welche Eigenschaften haben sie?

Sie haben damit einen Überblick darüber, ob die Methodik vollständig und verständlich dargestellt wurde. Im zweiten Schritt geht es darum, die Methodik zu bewerten. Ist die Methodik wirklich geeignet, die Fragestellung zu beantworten? Ganz wesentlich dafür sind drei Punkte:

- Was wurde eigentlich bei der UV variiert? Gibt das tatsächlich die Einflussgröße wieder, die untersucht werden sollte?
- Wie war der Versuchsplan? Ist wirklich sichergestellt, dass nur die UV eine Veränderung hervorrufen kann, oder gibt es unkontrollierte Störvariablen?
- Was wurde gemessen? Ist die AV eine gute Messung der Aspekte des Verhaltens und Erlebens, über die eine Aussage gemacht werden soll?

Überlegen Sie selbst, wie Sie diesen Versuch durchgeführt hätten. Was hätten Sie variiert, was gemessen? In welchem Versuchsplan? Warum haben die Forscher das anders gemacht? Diese Frage können auch helfen, die Methodik besser zu verstehen.

Schauen Sie sich jetzt den Methodikteil des Artikels genau an und beantworten Sie die oben dargestellten Fragen. Wir werden das in der nächsten Stunde gemeinsam besprechen.

7. Versuch Adaptationsniveau 1: Was ist der Reiz?

7.1. Der Gegenstand der Psychophysik

Sie haben inzwischen schon Einiges über die Methodik der Psychologie erfahren. Im Unterschied zur Physik gibt es in der Psychologie so viele Dimensionen, wie es unterschiedliche Erlebensweisen gibt. Etwas ist nur dann Gegenstand der Psychologie, wenn es erlebt wird. Die Metrik der Psychologie beschreibt, welche Dimensionen oder Qualitäten es gibt und welche Ausprägungen oder Quantitäten dabei erlebt werden.

Im ersten Versuch in diesem Buch haben Sie die Größe von Münzen geschätzt. Die Qualität, um die es dabei ging, war der Größeneindruck. Beim zweiten Versuch ging es um das Streckenmitteln. Die Qualität war hier die erlebte Streckenlänge, oder etwas genauer die Gleichheit von zwei Streckenlängen beim Einstellen der Streckenmitte.

Ihr Erleben konnten wir in beiden Versuchen nicht direkt messen, sondern mussten es aus dem Verhalten erschließen, wobei wir dieses wiederum physikalisch gemessen haben. Das ist eine Besonderheit der Wissenschaft „Psychologie“ – ihren Gegenstand kann man nicht direkt messen, sondern muss ihn indirekt erschließen. Aber auch bestimmte Teilbereiche anderer Wissenschaften haben dasselbe Problem. Denken Sie zum Beispiel an Elementarteilchen wie die Quarks. Man kann diese nicht direkt sichtbar machen, aber Bedingungen herstellen, wo diese Spuren hinterlassen. Aus diesen schließt man auf die Existenz der Teilchen. Ganz ähnlich ist es in der Psychologie, wenn man Verhalten misst, um dann auf ein bestimmtes Erleben zu schließen, dass für dieses Verhalten verantwortlich ist.

Im ersten Versuch war ihr Verhalten die Einstellung der Kreisgröße. Diese wurde dann durch das Programm ausgemessen. Der Messwert wurde als Beschreibung ihres Erlebens interpretiert. Im zweiten Versuch war das Verhalten die Teilung des Strichs. Gemessen wurde dann die Länge des Strichs, der für Sie genau der Mitte entsprach. Wir hatten dabei auch festgestellt, dass zu einem Erleben nicht nur ein physikalischer Messwert gehören kann, sondern dass dies auch mehrere sein können. Der Eindruck „Gleich“ gehörte zu einem Bereich von Messwerten, den wir Unsicherheitsintervall genannt haben. Das psychische „Gleich“ (ein Bereich von Werten) ist nicht identisch mit dem physikalischen „Gleich“ (ein Messwert). Das Erleben wurde aber mit Hilfe eines Maßes aus dem physikalischen Bereich beschrieben.

Dieses Vorgehen bezeichnet man als **Operationalisierung**: Da die Phänomene, die den Psychologen interessieren, häufig nicht direkt messbar sind, muss man sich überlegen, wie man sie messbar machen kann. Das subjektiv erlebte „Gleich“ wird beschrieben durch das Unsicherheitsintervall, das wiederum mit einer bestimmten Methode, z.B. dem Herstellungsverfahren, gemessen wurde. Dabei kann man sich natürlich immer fragen, ob das Messverfahren und der berechnete Messwert tatsächlich eine Operationalisierung des Erlebens sind. Ist die Standardabweichung der Messwerte eine gute Beschreibung für den Bereich, der für Sie „gleich“ war? Wäre es nicht besser, den ganzen Bereich, in dem Sie zumindest bei einem Reiz „gleich“ geurteilt haben, als Unsicherheitsintervall zu bezeichnen (also vom kleinsten zum größten Wert)? Bei der zweiten Definition würde jedes zufällige „Gleich“-Urteil zu einer Veränderung des Bereichs führen. Vielleicht ist deshalb ein Maß wie die Standardabweichung als Schätzung eines Streubereiches tatsächlich sinnvoller. Wir können hier diese Diskussion nicht weiter ausführen, aber Sie sollten sich merken, dass jede Operationalisierung willkürlich ist und begründet werden muss. Achten Sie bei dem nächsten Versuch einmal darauf, wie hier das Erleben gemessen wird. Ist das wirklich sinnvoll?

Aber diskutieren wir zunächst die Operationalisierung bei der Streckenmittelung noch etwas genauer. Abbildung 28 zeigt die Gegenüberstellung von physikalischen Werten und dem „Erleben“ für den Versuch des Streckenmitteln am Beispiel einer Versuchsperson. „Erleben“ habe

ich in Anführungszeichen gesetzt, weil die Grafik auch eine Operationalisierung ist, und natürlich nicht direkt das Erleben der Versuchsperson. Auf der x-Achse (Abszisse) sind die verschiedenen Strichlängen in Pixeln eingegeben, die der physikalischen Mitte entsprechen. Die Versuchsperson hatte jede Strecke mehrfach halbiert. Auf der y-Achse (Ordinate) sind die entsprechenden Ergebnisse eingetragen. Für jede Streckenlänge sieht man, welche Längen die Versuchsperson eingetragen hatte. Der Bereich, in dem diese Punkte liegen, ist der Bereich, in dem die Versuchsperson das „Gleich“ erlebt. Insofern sieht man in der Grafik das „Erleben“ als Streuung der Punkte. Wie gerade diskutiert, kann man jetzt überlegen, wie man das Erleben am besten quantifiziert. Wie kann man eine Zahl berechnen, die am besten das „Erleben“ wiedergibt?

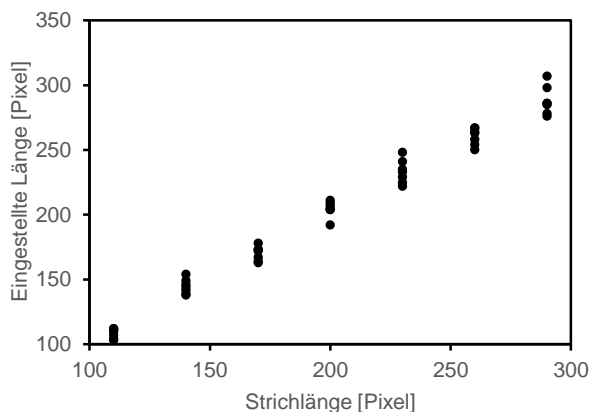


Abbildung 28: Ergebnisse der Streckenmittelung für eine Versuchsperson. Dargestellt sind die eingestellten Längen in Abhängigkeit von der Länge des Strichs bei der physikalischen Mitte.

Hilfreich ist es dazu, die Streuung der Einschätzung über zwei Geraden anzunähern (s. Abbildung 29). Man kann dann sehr gut das Weber'sche Gesetz erkennen. Je länger die Striche sind, umso größer wird die Streuung der Werte. Man kann jetzt überlegen, mit welchem Messwert man diese Eigenschaft des Erlebens am besten wiedergeben kann. Im Versuch hatten wir die Standardabweichung pro Streckenlänge berechnet. Wenn wir für die Versuchsperson in Abbildung 29 für jede Streckenlänge die Differenz aus größtem und kleinstem Wert berechnen würden, würde diese auch sehr gut diese Eigenschaft wiedergeben. Es gibt also häufig verschiedene Methoden, um das Erleben indirekt zu messen.

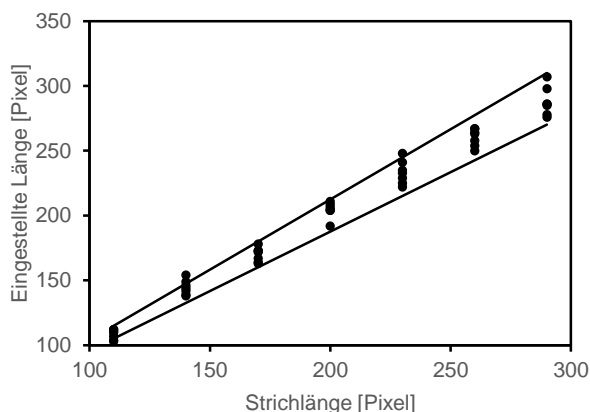


Abbildung 29: Verdeutlichung des Weber'schen Gesetzes für die Daten aus Abbildung 28.

Zurück zur Bedeutung dieses Ergebnisses: Während die Physik hier additive Zusammenhänge aufweist, findet sich demnach im Erleben eine prozentuale Metrik. Das Unsicherheitsintervall

ist eben nicht für alle Strichlängen gleich, sondern prozentual abhängig von der Länge der Striche. Um dieses Gesetz nachzuweisen, wurde die Strichlänge als unabhängige Variable (UV) variiert und das Erleben (eingestellte Länge) als abhängige Variable (AV) gemessen. In der Grafik ist also die UV auf der x-Achse dargestellt, die AV auf der y-Achse.

Die Abbildung zeigt also genau den Zusammenhang zwischen UV und AV als Funktion, also

$$AV = f(UV)$$

Diese Funktion gibt für unsere Versuche wieder, wie physikalische Objekte erlebt werden. Sie ist also eine Art Übersetzungsregel vom Physikalischen ins Psychische. Bei der Untersuchung dieser Übersetzungsregeln wurden in den frühen Zeiten der Psychologie drei Dinge festgestellt:

- 1) Man kann die Übersetzung sehr gut beschreiben. In einem mittleren Wertebereich findet man häufig lineare Zusammenhänge.
- 2) Bestimmte Reize werden überhaupt nicht wahrgenommen. Dies führte zur Bestimmung von Schwellen.
- 3) Die Wahrnehmung hat eine bestimmte Messgenauigkeit. Die Bestimmung dieser Messgenauigkeit führte zum Weber'schen Gesetz.

Dieses Vorgehen – Übersetzungsregeln von der Physik in die Psychologie zu finden - hat diesem Bereich der Psychologie auch seinen Namen gegeben: Psychophysik – die Frage, wie sich physikalische Phänomene im Erleben, dem Psychischen, abbilden.

Dieses Vorgehen stößt allerdings schnell an seine Grenzen, wie Abbildung 30 zeigt. Der graue Kreis links wirkt deutlich heller als der graue Kreis rechts. Wenn man allerdings die Seite faltet, sodass beide Kreise direkt nebeneinander liegen, sieht man plötzlich, dass sie genau gleich hell sind. Die Ursache dafür ist der Kontext. Das menschliche Erleben von Helligkeit ist eben nicht einfach eine Funktion der physikalischen messbaren Helligkeit, sondern bezieht die Helligkeit der Umgebung mit ein. Dasselbe Grau wirkt vor einem schwarzen Hintergrund relativ hell, vor einem weißen Hintergrund relativ dunkel. Eine Maßfunktion, die die physikalische Helligkeit auf dieses Erleben direkt abbildet, existiert nicht. Oder:

$$AV \neq f(UV)$$

Die Ursache dafür liegt darin, dass der Reiz, den der Mensch erlebt, nicht nur der graue Kreis ist, sondern der graue Kreis in Relation zu einem schwarzen oder weißen Hintergrund. Damit ergibt sich eine ganz neue Frage:

- Was ist eigentlich der Reiz, der einem bestimmten Erleben zugrunde liegt?

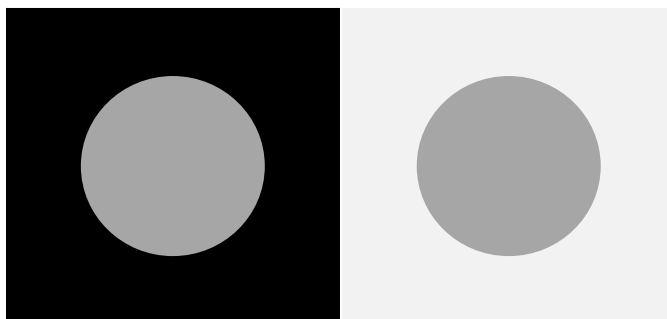


Abbildung 30: Kontextabhängigkeit der Wahrnehmung. Zur weiteren Erklärung s. Text.

Dieser Frage waren wir in etwas veränderter Form im Rahmen dieses Buchs bereits begegnet, als wir uns mit der Instruktion beschäftigt hatten. Dort hatte sich gezeigt, dass je nach Definition des „Gleich“ ganz unterschiedliche Urteile auftraten. Wenn man nur dann „gleich“ urteilt, wenn man trotz größter Anstrengung keinen Unterschied finden kann, ist das Unsicherheitsintervall

sehr eng. Wenn man solange „gleich“ urteilt, wie man sich nicht ganz sicher ist, dass ein Unterschied vorliegt, wird das Unsicherheitsintervall sehr breit. Hier waren durch die Instruktion unterschiedliche „Reize“ erzeugt worden, die dann zu unterschiedlichen Urteilen führten.

Wir hatten bei der Instruktion zu den Versuchen auch immer darauf hingewiesen, dass man möglichst konzentriert und ungestört arbeiten sollte. Man sollte auch die Urteile immer treffen, ohne nachzumessen, ohne sich an der Größe des Bildschirms zu orientieren usw. Damit war immer versucht worden, den Kontext der Urteile möglichst konstant zu halten. Wie die Abbildung 30 zeigt, wird dies aber eigentlich dem „wahren“ Leben nicht gerecht: Jeder Reiz, den wir erleben, ist in einen Kontext eingebettet. Wenn der Kontext tatsächlich so sehr das Erleben beeinflusst, dann muss die Untersuchung von Kontexteffekten ein ganz zentraler Bereich der Psychologie sein. Das ist auch so! Dieser Bereich ist unter der Bezeichnung „Bezugssystemtheorie“ bekannt geworden. Der Kontext wirkt also als Bezugssystem. Was dies genau bedeutet, werden wir mit Hilfe des nächsten Versuchs demonstrieren. Die zentrale Frage ist hier:

- Wie ändert sich das Urteil, wenn sich der Kontext verändert?

Im Schema des Experiments bedeutet dies, dass man den Kontext als unabhängige Variable variiert und die Veränderung der Urteile als abhängige Variable misst. Auch dies werden Sie wieder im Selbstversuch untersuchen. Dabei ist es wichtig, dass Sie jetzt gleich drei Versuche machen, den ersten Versuch für dieses Kapitel und dann die beiden Versuche für das nächste Kapitel, aber ohne dieses vorher zu lesen. Wir werden gleich erklären, warum das wichtig ist.

7.2. Instruktion Adaptationsniveau I

In dem Versuch soll es insgesamt darum gehen, wie sich Urteile verändern, wenn sich der Kontext verändert. Um dies zu untersuchen, werden wir zunächst verschiedene Reize vorgeben und beurteilen lassen. In dem zweiten Versuchsteil werden wir dann den Kontext variieren und dann die Auswirkungen untersuchen.

Als Dimension verwenden wir in diesem Versuch wieder die Tonhöhe. Diesmal werden Sie allerdings eine andere Bandbreite von Tönen hören als im letzten Versuch. Dadurch können wir den Einfluss des Kontexts für eine größere Menge von Reizen untersuchen. Sie werden dabei auch eine andere Skalierungsmethode kennenlernen. Sie sollen nach jedem Ton die Tonhöhe direkt auf folgender Skala beurteilen:

Sehr sehr hoch
Sehr hoch
Hoch
Mittel hoch
Mittel
Mittel tief
Tief
Sehr tief
Sehr sehr tief

Beim Versuch wird diese Skala immer dann gezeigt, wenn Sie ein Urteil abgeben sollen.

Damit Sie einen Eindruck bekommen, welche Töne im Versuch vorkommen, werden Ihnen die 15 Töne zunächst alle nacheinander in zufälliger Reihenfolge vorgespielt. Hören Sie sich die

Töne bitte konzentriert an! Danach geht der eigentliche Versuch los. Auch hier werden wieder die 15 Töne in zufälliger Reihenfolge vorgeben. Bei jedem Durchgang hören Sie zwei Töne. Der erste Ton ist das „Achtung“-Signal. Hören Sie sich dann den zweiten Ton danach genau an und geben Sie dazu Ihr Urteil ab. Am Ende des Versuchs wird eine Tabelle mit Ihren Urteilen gezeigt. Diese können Sie wieder kopieren und in Excel weiterbearbeiten. Die Tonserie wird dreifach dargeboten, um auf diese Weise stabilere Urteile von Ihnen zu erhalten. Insgesamt sollen Sie 45 Töne beurteilen.

Bitte führen Sie diesen Versuch in einer ruhigen Umgebung durch, in der Sie ungestört arbeiten können. Stellen Sie die Lautstärke vorher so ein, dass die Töne gut hörbar und angenehm sind. Konzentrieren Sie sich dann und beginnen Sie den Versuch! Am Ende des Versuchs wird wieder eine Tabelle mit den Ergebnissen angezeigt, die Sie nach Excel kopieren können. Speichern Sie die Ergebnisse dort und führen Sie dann die nächsten beiden Versuche durch!

Und hier geht es zum Versuch:

- <https://cgi-psych.rz.tu-bs.de>

7.3. Instruktion Adaptationsniveau IIa und IIb

In diesem zweiten Versuchsteil geht es jetzt um den Einfluss des Kontexts. Diesen werden wir auf eine sehr einfache Art verändern: Der „Achtung“-Ton wird einmal sehr hoch, das andere Mal sehr tief sein. Es geht also darum, ob ein solcher Ton, der direkt vor dem eigentlich zu beurteilenden Ton gehört wird, das Erleben verändert. Man bezeichnet solche Reize als „Anker“.

Wichtig für den Versuch ist, dass Sie nicht den Anker-Ton beurteilen, sondern den Ton direkt danach. Denken Sie einfach daran, dass der erste Ton ein „Achtung“-Signal ist und konzentrieren Sie sich auf den zweiten Ton und beurteilen Sie diesen.

In jedem der beiden Versuchsteile IIa und IIb erhalten Sie wieder dieselbe Serie wie im Versuch I. Auch hier wird die Serie dreifach wiederholt, um ein stabiles Urteil zu erhalten. Bitte versuchen Sie **nicht**, sich an Ihre Urteile aus dem ersten Versuch zu erinnern, sondern beurteilen Sie jeden Ton für sich mit Hilfe der Skala. Geben Sie einfach an, wie Sie den Ton im Hier und Jetzt erleben.

Bitte führen Sie diesen Versuch in einer ruhigen Umgebung durch, in der Sie ungestört arbeiten können. Stellen Sie die Lautstärke vorher so ein, dass die Töne gut hörbar und angenehm sind. Konzentrieren Sie sich dann und beginnen Sie den Versuch! Machen Sie bitte eine kurze Pause zwischen dem Versuchsteil IIa und IIb, damit Sie sich gut konzentrieren können. Übertragen Sie nach jedem Versuch die Tabelle, die Ihnen das Programm zeigt, nach Excel.

Und hier geht es zum Versuch:

- <https://cgi-psych.rz.tu-bs.de>

7.4. Auswertung Versuch Adaptationsniveau I

Wir haben Sie bei diesem Versuch nicht vollständig instruiert. Es waren am Anfang 15 Töne vorgespielt worden, um Sie über die möglichen Reize zu informieren. Allerdings handelte es sich dabei nur um 5 unterschiedliche Töne, von denen jeder dreimal vorgespielt wurde. Haben Sie dies gemerkt? Die meisten Probanden haben kein absolutes Gehör und merken dies nicht. Wir haben Sie bewusst nicht darüber informiert, weil Sie sonst vermutlich die Antwortskala mit Ihren neun Kategorien nicht ausgenutzt hätten.

Bei dem Versuch haben Sie dann auch in jedem Durchgang die fünf unterschiedlich hohen Töne gehört. Da Sie insgesamt neun Durchgänge durchgeführt haben, haben Sie also pro Ton 9 Urteile. Um Ihr typisches Urteil für jeden Ton zu erhalten, berechnet man den Mittelwert über diese 9 Urteile. Dazu müssen Sie im ersten Schritt Ihre Urteile nach Excel kopieren. Die Ausgabe des Programms für Excel ist schon so vorsortiert, dass die gleichen Töne untereinander stehen. Als Überschrift finden Sie die Tonhöhe in Herz. Wir haben Töne von 750, 775, 800, 825 und 850 Hz verwendet.

Pro Spalte, also für jeden Ton, berechnen Sie dann den Mittelwert. Erstellen Sie jetzt eine Grafik aus Ihren Daten, bei der Sie die unabhängige Variable (Tonhöhe) auf der x-Achse und die abhängige Variable (Urteil) auf der y-Achse eintragen. Es sollte sich annähernd eine lineare Beziehung ergeben, wie es die Abbildung 31 für eine Gruppe von Studenten zeigt. In dieser Grafik wurde für die y-Achse nur der Ausschnitt der möglichen Urteile gewählt, der auch verwendet wurde. Auf diese Weise ist der Zusammenhang zwischen UV und AV besser zu sehen. Sie erkennen aber auch, dass die Urteile bei diesen Studenten im mittleren bis hohen Bereich lagen. „4“ entsprach „mittel tief“, „8“ sehr hoch.

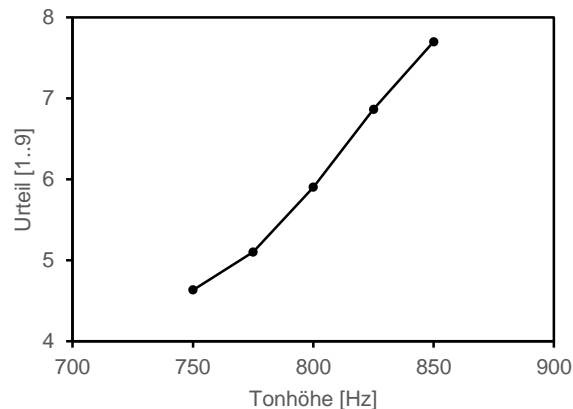


Abbildung 31: Mittelwert der Urteile von einer Gruppe von Studenten ($n = 23$) über die Tonhöhe.

Man muss sich einmal klarmachen, was diese Gerade bedeutet. Als erstes zeigt sich, dass die Reihenfolge oder Ordnung der Töne (von hoch zu tief) in derselben Weise auch in den Urteilen zu finden ist. Je höher ein Ton, desto größer auch das Urteil. Diese Systematik würde aber auch gelten, wenn es keine Gerade ist, sondern z.B. ein Zusammenhang wie in Abbildung 32.

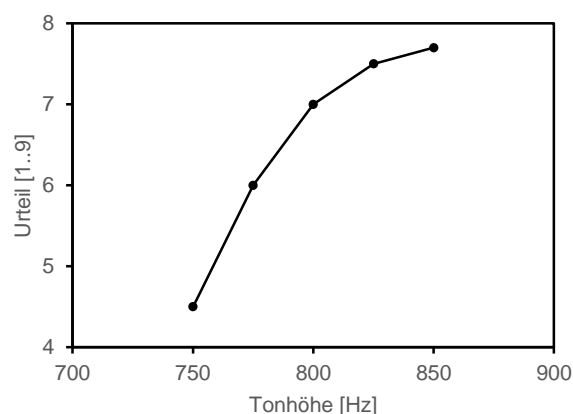


Abbildung 32: Beispiel für einen (fiktiven) nicht-linearen Zusammenhang zwischen Tonhöhe und Urteil.

Auch für die Daten in dieser Grafik gilt: Je höher der Ton, desto größer das Urteil. Wenn ein linearer Zusammenhang vorliegt, kommt eine weitere Interpretation hinzu: Ein bestimmter Zuwachs der Frequenz des Tons entspricht einem bestimmten Zuwachs des Urteils. In der Abbildung 31 findet sich bei einem Zuwachs von 25 Hz ein Urteilszuwachs von ungefähr 0.7.

Werten Sie jetzt Ihre Daten entsprechend aus. Schauen sie sich zunächst Ihre eigene Abbildung an. Ergibt sich auch bei Ihnen annähernd eine Gerade?

Berechnen Sie außerdem die Differenzen der Urteile zwischen den jeweils benachbarten Frequenzen, wie es in Tabelle 10 vorgegeben ist. Wenn zwischen den Frequenzen und Ihren Urteilen eine lineare Beziehung besteht, so sollten sich diese Differenzen sehr ähnlich sein.

Tabelle 10: Tragen Sie hier die Differenzen zwischen Ihren Urteilen an. Sie sollten annähernd gleich sein.

Urteil(775 Hz) – Urteil(750 Hz) =	
Urteil(800 Hz) – Urteil(775 Hz) =	
Urteil(825 Hz) – Urteil(800 Hz) =	
Urteil(850 Hz) – Urteil(825 Hz) =	

Diese Untersuchung des Zusammenhangs zwischen physikalischen Reizen und dem Erleben ist typisch für die klassische Psychophysik. In welcher Beziehung stehen die messbare Reizausprägung und die Ausprägung im Erleben? In dem Versuch hatten wir eine lineare Beziehung gefunden. Wenn wir hier eine Regressionsgerade anpassen würden, könnten wir für jede Reizgröße das entsprechende Urteil vorhersagen. Wir könnten also die Beziehung zwischen UV (Tonhöhe in Hz) und AV (Urteil) als mathematische Gleichung beschreiben.

7.5. Exkurs: Psychologische Bedeutung von Zahlen

In dem Versuch hatten Sie Ihr Erleben mit Hilfe von Kategorien beschrieben. Der Ton war „sehr hoch“ oder „mittel hoch“. Für jeden Ton lagen am Ende neun derartige Urteile vor. Was ist jetzt Ihr typisches Urteil? Im Abschnitt oben und für die Abbildung 32 hatten wir einfach jeder Kategorie eine Zahl zugeordnet. „Sehr hoch“ wurde so zu einer 8, „mittel hoch“ zu einer 6. Aus diesen Zahlen konnten Sie (oder Excel) dann sehr einfach einen Mittelwert berechnen. Aber macht das überhaupt Sinn? Ist „sehr hoch“ in Ihrem Kopf eine 8?

Wir Menschen denken meistens nicht in Zahlen. Aber wir haben gelernt, die Ausprägung unseres Erlebens (die Quantität) sprachlich auszudrücken. Es war für Sie im Versuch überhaupt kein Problem, einen Ton als „sehr hoch“ zu beschreiben. Die Frage, die wir uns hier wissenschaftlich stellen, ist allerdings mit Hilfe von derartigen Kategorien schwer zu beantworten. Wie gut entspricht das menschliche Erleben von Tonhöhe den physikalischen Eigenschaften? Wie genau ist das menschliche Erleben?

Die physikalische Eigenschaft messen wir, indem wir diesen Eigenschaften auf eine genau definierte Weise Zahlen zuordnen. Der Urmeter ist hier ein gutes Beispiel: Die Länge eines Strichs wird gemessen, indem wir den Strich neben das Urmeter (oder eine Kopie davon) legen und dort ablesen, welcher Zahlenwert dort steht. Die Zahlen wurden willkürlich definiert. Um jetzt die physikalische Messung mit dem Erleben zu vergleichen, ist es hilfreich, wenn das Erleben auch in Zahlen ausgedrückt wird. Auch dies ist willkürlich: Wir haben entschieden, dass „sehr hoch“ als 8 codiert wird. Wenn wir dies tun, können wir Mittelwerte aus den Urteilen berechnen und dann die Beziehung zwischen physikalischen Eigenschaften und dem Erleben grafisch darstellen wie in Abbildung 32. Oder wir können eine Funktion berechnen, die die Beziehung zwischen der Tonhöhe in Hz und dem Erleben zum Beispiel als lineare Gleichung beschreibt. Bei unserer Gruppe von Studenten ergibt dann:

$$\text{Erlebte Tonhöhe [1..9]} = 0.0316 * \text{Tonhöhe [Hz]} - 19.2$$

Ein Ton von 800 Hz würde nach dieser Formel zu einer erlebten Tonhöhe von 6.1 führen (man kann solche Berechnungen übrigens sehr einfach in Excel durchführen, wenn Sie das nachrechnen wollen). Dies passt sehr gut zu den tatsächlichen Urteilen aus der Abbildung. Man kann dann auch wieder umcodieren in die verwendeten Kategorien. Die 6 hatte der Kategorie „mittel hoch“ entsprochen. Ein Ton von 800 Hz wird also als „mittel hoch“ erlebt.

Das einzige Problem bei dieser Art des Vorgehens ist, dass Zahlen eine ganze Reihe von Eigenschaften haben, die wir so gut gelernt haben, dass wir ganz unwillkürlich an diese Eigenschaften denken, wenn wir mit Zahlen umgehen. Eine 6 ist doppelt so groß wie eine 3, oder? Aber ist ein Ton, der als „mittel hoch“ (von uns als 6 codiert) erlebt wird, doppelt so hoch wie ein Ton, der als „tief“ (von uns als 3 codiert) beurteilt wird? Ist „mittel hoch“ doppelt so hoch wie „tief“? Das kann man auf der Basis unseres Versuchs nicht beantworten. Man müsste dazu anders vorgehen. Man könnte zum Beispiel einen Ton vorspielen und der Versuchsperson die Aufgabe geben, einen zweiten Ton so einzustellen, dass er halb so hoch ist wie der erste Ton. Ein Herstellungsverfahren, bei dem die Versuchsperson selbst die Tonhöhe einstellen kann, wäre dazu eine gute Methode (wenn Sie möchten, können Sie mit dem Versuch „Größeneindruck“ für die Größe von Kreisen ausprobieren, wie gut Sie Kreisgrößen halbieren können). So wie wir den Versuch mit den Tonhöhen hier durchgeführt haben, darf man vermutlich diese Eigenschaft der Zahlen (Verhältnisse zu bilden) nicht interpretieren.

Welche Eigenschaften darf man dann bei diesem Versuch interpretieren? Welche Eigenschaften von Zahlen gibt es überhaupt?

Die fundamentale Eigenschaft von Zahlen ist „gleich/ungleich“. Zwei Reize, denen die gleiche Zahl zugeordnet wird, werden als gleich erlebt. Reize mit unterschiedlichen Zahlen sind auch im Erleben unterscheidbar. Dies entspricht auch dem Grundgedanken von Kategorien. Wenn der eine Ton als „mittel hoch“, der andere als „tief“ erlebt wird, unterscheidet sich das Erleben. Man sollte diesen beiden Kategorien also auch unterschiedliche Zahlen zuordnen. Wenn man sich darauf einigt, dass man nur die Eigenschaft „gleich/ungleich“ von Zahlen interpretiert, dann könnte man aber auch die Kategorie „mittel hoch“ mit 3 und die Kategorie „tief“ mit 6 codieren.

Bei unseren Kategorien war aber neben dem „Gleich/Ungleich“ noch eine zweite Eigenschaft enthalten, die Ordnung von Reizen. Die Reihenfolge der einzelnen Kategorien entspricht einer zunehmenden oder abnehmenden Stärke des Erlebens. Ein „sehr tiefer“ Ton ist tiefer als ein „mittlerer“ Ton. Also sollte auch die Zahl für die Kategorie „sehr tief“ kleiner sein als die Zahl für die Kategorie „mittel“. Ob man dafür die Zahlen 2 „sehr tief“ und 5 „mittel“ oder 1 „sehr tief“ und 100 „mittel“ verwendet, ist egal, solange man nur die Ordnung interpretiert.

Bei der Interpretation der Abbildung 32 hatten wir noch eine weitere Eigenschaft gesehen. Wenn der Reiz um einen bestimmten Betrag verändert wird, ändert sich das Erleben ebenfalls jeweils um einen bestimmten Betrag. 25 Hz entsprachen bei unseren Versuchspersonen ungefähr 0.7 bei unseren codierten Kategorien. Ein kleiner Unterschied oder eine kleine Differenz zwischen Zahlen entspricht auch einem kleinen Unterschied im Erleben. Zwei Reize, die sich um 50 Hz unterscheiden, liegen auch im Erleben der Tonhöhe weiter auseinander als zwei Reize, die sich nur um 25 Hz unterscheiden.

Bei den von uns verwendeten Kategorien macht es demnach vermutlich Sinn, die Zahlen so zu wählen, dass der Unterschied zwischen zwei Kategorie bei den verwendeten Zahlen jeweils gleich ist. So wie wir oben die Zahlen 1 bis 9 aufsteigend den Kategorien zugeordnet hatten, war genau das geschehen. Zwei benachbarte Kategorien unterscheiden sich jeweils um „1“.

Man hätte auch [2, 4, 6, 8, 10, 12, 14, 16, 18] oder [5..13] wählen können. Auch in diesen Zahlenreihen ist der Abstand zwischen allen zwei benachbarten Kategorien jeweils gleich groß.

Schwierig ist die vierte wesentliche Eigenschaft von Zahlen, das Bilden von Verhältnissen. Physikalisch sind 1000 Hz doppelt so hoch wie 500 Hz. Aber gilt das auch für die entsprechenden Kategorien? Wie am Anfang dargestellt wurde, ist das bei der Art, wie wir die Töne beurteilt hatten, wahrscheinlich nicht der Fall. Mit einem Herstellungsverfahren, bei dem die Versuchsperson Verhältnisse von Tonhöhen herstellt (z.B. „A ist doppelt so hoch wie B“ oder „C ist halb so hoch wie B“) könnte man entsprechende Zahlen verwenden ($B = 2$, $A = 4$, $C = 1$). Bei diesen darf man dann auch die Verhältnisse interpretieren (1 ist halb so groß wie 2, also ist C halb so hoch wie B).

Um für den Leser schnell und klar deutlich zu machen, welche Eigenschaften von gemessenen Zahlen man interpretieren darf, beschreibt man dies als „Messung auf einem bestimmten Skalenniveau“.

- „Gleich/ungleich“ entspricht einer Messung auf dem Nominalniveau (Nominalskala)
- „Größer/kleiner“ entspricht einer Messung auf dem Ordinalniveau (Ordinalskala)
- Der Vergleich von Abständen entspricht einer Messung auf dem Intervallniveau (Intervallskala)
- Ist das Bilden von Verhältnissen sinnvoll, spricht man von einer Messung auf dem Verhältnissniveau (Verhältnisskala)

Der Vollständigkeit halber soll noch die letzte Art von Skala, die Absolutskala, erwähnt werden. Hier hat die Zahl für sich eine bestimmte Bedeutung. Dies ist z.B. dann der Fall, wenn man die Anzahl von Objekten zählt. Jede Zahl steht hier genau für eine ganz bestimmte Menge von Objekten.

Wir hatten uns mit diesen unterschiedlichen Eigenschaften von Zahlen beschäftigt, da wir zur Beurteilung der Tonhöhe Kategorien verwendet hatten, denen wir sinnvoll Zahlen zuordnen wollten, um die Beziehung zwischen physikalischen Eigenschaften (Tonhöhe in Hz) und erlebter Eigenschaft („sehr laut“ usw.) grafisch und als Funktion darzustellen. Wir hatten gesehen, dass wir diese Zahlen auf Intervallniveau interpretieren können. Dies wurde in Abbildung 32 deutlich, bei der sich ein linearer Zusammenhang zwischen physikalischer und erlebter Tonhöhe zeigte. Gleiche physikalische Abstände zwischen Tönen entsprechen gleichen erlebten Unterschieden in der Tonhöhe. Im nächsten Abschnitt wird dieses Ergebnis weiter diskutiert.

7.6. Absoluturteile und Adaptationsniveau

Man kann diese Daten aber auch anders auswerten und ganz im Psychischen, im Erleben, bleiben. In Abbildung 33 sind die mittleren Urteile zu den fünf Tönen über der verwendeten Skala aufgetragen. So wurde der Ton mit 850 Hz z.B. als „8 = sehr hoch“ beurteilt. Der tiefste verwendete Ton von 750 Hz wurde zwischen „4 = mittel tief“ und „5 = mittel“ eingeordnet. Insgesamt lag damit unsere Tonreihe im Bereich zwischen mittel tief und sehr hoch, also eher bei höheren Tönen.



Abbildung 33: Die Ergebnisse aus dem Studentenversuch in Abbildung 31 aufgetragen über der Skala, die verwendet wurde. Die einzelnen Kreuze sind die Mittelwerte der Urteile bei den fünf Tönen.

Mit dieser Darstellung haben wir eine rein psychologische Auswertung und Darstellung durchgeführt. Die physikalischen Messwerte sind hier eigentlich nur die „Namen“ oder „Labels“ für die 5 verschiedenen Töne, die verwendet wurden. Die eigentliche Aussage wäre hier: „Der höchste verwendete Ton war ein sehr hoher Ton (Skalenwert 8), der niedrigste ein mittel tief bis mittlerer Ton (Skalenwert 4.5).“

Wie ist diese Aussage entstanden? In allen früheren Versuchen hatten wir immer Reize verglichen. Ein Reiz war kleiner, genauso groß oder größer als ein anderer Reiz (Münzgrößen). Ein Strich war länger, kürzer oder genauso lang wie ein anderer (Streckenmitteln). Ein Ton war höher oder tiefer oder gleich hoch wie ein Vergleichston (Grenzverfahren). Hier hatten also immer Vergleiche stattgefunden. Entsprechend nennt man diese Art von Urteilen auch „Vergleichsurteile“.

In dem Versuch jetzt hatten Sie dagegen immer nur einen Ton beurteilt. Sie sollten jeden Ton für sich auf der Skala einordnen. Sie haben also nicht verglichen, sondern jeden Ton „absolut“ beurteilt. Solche Arten von Urteilen nennt man deshalb „Absoluturteile“. Aber wie ist ein solches Urteil überhaupt möglich? Woher weiß man, was ein „sehr sehr hoher Ton“ ist? Dieses Urteil macht nur dann einen Sinn, wenn man auf ein Vorwissen zurückgreift, also weiß, was typische hohe und tiefe Töne sind. Dieses Vorwissen wird in der Psychophysik als „Bezugssystem“ bezeichnet.

- Bezugssystem: Das Vorwissen über die möglichen Ausprägungen einer bestimmten Dimension, das für absolute Urteile genutzt wird.

In dem Versuch haben wir versucht, ein solches Bezugssystem herzustellen, indem wir Ihnen vor dem Versuch die Töne, die Sie dann beurteilen sollten, vorgespielt haben. Wenn Sie nur diese Menge von Tönen im Kopf gehabt hätten, dann hätten Sie den niedrigsten Ton als „sehr sehr tief“ und den höchsten als „sehr sehr hoch“ beurteilen müssen. Schauen Sie sich Ihre Werte und Abbildung 33 an. In der Abbildung wurde der tiefste Ton als „mittel tief bis mittel“ beurteilt, der höchste als „sehr hoch“. Was kann man darauf schließen? Offensichtlich hatten Sie neben dem durch das Programm vorgespielten Bezugssystem (die 15 Töne, die Sie vorher gehört haben) noch ein anderes Bezugssystem in Ihrem Kopf. Kein Wunder, schließlich hat jeder von Ihnen im Laufe seines Lebens bereits eine Vielzahl von Tönen gehört und sich diese Bandbreite gemerkt. Deshalb weiß natürlich jeder, dass es noch viel tiefere Töne gibt als die, die Ihnen hier vorgespielt wurden.

Aus diesen Überlegungen sind also folgende Punkte festzuhalten:

- Ein Absoluturteil ist nur möglich, wenn man auf ein Bezugssystem zurückgreifen kann.

- Bezugssysteme werden erlernt, indem man erlebt, welche Ausprägungen einer Erlebensdimension (Tonhöhe, Strichlänge usw.) möglich sind.
- Für jede Dimension kann es viele unterschiedliche Bezugssysteme geben.

Um den letzten Punkt noch weiter zu verdeutlichen: Statt mit Sinustönen hätte man mit den Tönen einer Flöte oder eines Basses arbeiten können. Beide Instrumente haben sehr unterschiedliche Bereiche von Tönen, die damit gespielt werden können. Die meisten Menschen kennen diese Bereiche einigermaßen gut. Wenn man jetzt dieselbe mittlere Frequenz einmal mit dem Bass und einmal mit der Flöte spielt und dann Absoluturteile abfragt, so wäre diese Frequenz bei einer Flöte vermutlich relativ tief, bei dem Bass relativ hoch. Man sieht hier direkt die Bedeutung von Bezugssystemen: Derselbe Reiz wird ganz automatisch unterschiedlich beurteilt, wenn man sich auf verschiedene Bezugssysteme bezieht.

Damit haben wir bei unserem Versuch „Adaptationsniveau“ einen Versuchsfehler begangen. Es wurde vor dem Versuch nicht eindeutig genug instruiert, auf welches Bezugssystem Sie sich bei Ihren Urteilen beziehen sollen. Man hätte also vielleicht instruieren sollen:

„Vergessen Sie alle Töne, die Sie jemals gehört haben! Es geht nur um das Hier und Jetzt! Wir spielen Ihnen jetzt 15 Töne aus einem bestimmten Frequenzbereich vor. Für diesen Bereich gibt es sehr sehr tiefe und sehr sehr hohe Töne. Versuchen Sie, sich diesen Frequenzbereich möglichst gut einzuprägen. Danach bekommen Sie die Töne einzeln zu hören. Denken Sie an den Frequenzbereich und beurteilen Sie dann die Töne hinsichtlich ihrer Lage in diesem Bereich. Verwenden Sie dazu die Kategorien von sehr sehr tief bis sehr sehr hoch.“

Bei dieser Instruktion wären die Mittelwerte völlig anders ausgefallen und hätten vermutlich den Bereich von 1 bis 9 sehr gut abgedeckt. Aber hat dieses Vorgehen, haben diese Urteile noch etwas mit Ihrem Erleben zu tun? Man gibt doch eigentlich die Urteile schon durch die Instruktion vor. Eigentlich geht es dann nur noch darum, wie genau, wie zuverlässig Ihre Urteile sind. Aber das ist natürlich auch eine wichtige Frage. Wir hatten ja oben gesehen, dass die Frage der Unterscheidungsfähigkeit, der Genauigkeit des Erlebens, eine zentrale Eigenschaft der Metrik des Psychischen ist. Wenn es also nur um die Genauigkeit des Erlebens geht, wäre diese Instruktion also sehr sinnvoll gewesen.

Diese Kritik stimmt auch aus einem anderen Grund nur teilweise. Man gibt ja nur die Pole, die Eckpunkte, den Bereich vor, in dem geurteilt werden soll. In der Physik nennt man dies Eichung. Bei dem Thermometer wird festgesetzt, wo der Nullpunkt liegt. Der Unterschied zwischen Fahrenheit und Grad Celsius zeigt, dass diese Festlegung willkürlich ist. Dennoch ist die Gradangabe sehr sinnvoll. Man muss eben nur wissen, ob hier Grad Celsius oder Fahrenheit gemeint sind. Genauso sind auch die Urteile tief, mittel und hoch sehr sinnvoll. Man muss eben auch hier nur wissen, welches Bezugssystem der Mensch im Kopf gehabt hat, als er seine Urteile getroffen hat. Und wenn man wissen möchte, wie gut Menschen in der Lage sind, Unterschiede zwischen verschiedenen Frequenzen auch entsprechend zu erleben und zu beurteilen, dann macht die Verankerung die Urteile verschiedener Menschen vergleichbar, aber gibt nicht das Ergebnis bereits vor. Offen bleibt auch, wie die Beziehung im Bereich zwischen den Polen aussieht, ob sie z.B. linear wie in Abbildung 31 ist.

Absoluturteile machen also nur dann Sinn und sind nur dann interpretierbar, wenn man das Bezugssystem kennt. Eigentlich ist der Begriff „Absoluturteil“ damit auch etwas irreführend. Es findet ja doch ein Vergleich statt, allerdings nicht mit anderen Reizen, sondern mit dem Wissen, welche Ausprägungen der Reize überhaupt in diesem Kontext vorkommen können, in dem das Urteil gefällt werden soll. Der Begriff ist sinnvoll, um sich von einer Situation abzugrenzen, in der man Reize miteinander vergleichen soll. Außerdem hat er sich etabliert, sodass

wir ihn im Folgenden auch beibehalten werden. Sie sollten sich aber merken, dass das Absoluturteil eigentlich auch relativ ist.

Die Basis der Absoluturteile ist das Bezugssystem. Welche Struktur hat ein solches System und wie bildet es sich aus? Wie sieht das in unserem Kopf aus? Natürlich kann man diese Frage nicht direkt beantworten, also im Kopf nachsehen, was dort gespeichert ist. Wir müssen uns wie in allen Bereichen der Psychologie der Antwort indirekt nähern. Dazu überlegt man, wie so etwas gespeichert sein könnte und wie sich das wiederum auf die Urteile auswirken müsste. Dann kann man experimentell prüfen, ob die Ergebnisse diese Überlegungen bestätigen.

Wir hatten diskutiert, dass ein Bezugssystem die Absoluturteile verankert, d.h. die Pole oder die extremen Ausprägungen enthält. Damit ist der Bereich klar, in dem die Ausprägungen liegen. Die einfachste Art eines Bezugssystems wäre damit eines, das nur die Repräsentation der beiden Pole enthält. Man müsste dann bei jedem Reiz, den man hört, vergleichen, ob der näher am einen oder anderen Pol liegt, also z.B. eher hoch oder eher tief ist. So ein System würde allerdings nicht ausreichen, um die Einordnung in die verschiedenen Kategorien (z.B. „sehr hoch“) zu ermöglichen.

Ganz am Anfang der Übung hatten wir auch die besondere Rolle der Mitte beschrieben. Wenn man Erfahrungen mit Reizen sammelt, bildet sich ein Bereich aus, der weder zum einen noch zum anderen Pol gehört, also weder groß noch klein, weder hoch noch tief ist. Er wird mit dem Urteil „mittel“ belegt. Zu den Polen kommt so ein dritter Punkt hinzu. Mit zunehmender Erfahrung kann sich der „Mittel“-Bereich als eigene Kategorie stabilisieren und sich dann wiederum zwischen „Mittel“ und „Groß“ bzw. „Mittel“ und „Klein“ ein neuer Bereich bilden, der zu keiner dieser beiden Kategorien gehört. Auf diese Weise könnte ein Bezugssystem mit zunehmender Erfahrung immer feiner werden, sich also immer mehr gut unterscheidbare Kategorien bilden. Man würde sich also vorstellen, dass in einem Bezugssystem sowohl die beiden Pole repräsentiert sind als auch verschiedene unterscheidbare Ausprägungen dazwischen.

Der erste Schritt in diesem Prozess der Bezugssystementstehung ist die Bildung der „Mitte“. Sie spielt eine ganz besondere Rolle und gibt die Lage des Bezugssystems an. Man gewöhnt sich daran, dass die Reize in einem bestimmten Bereich liegen, beim Bass also insgesamt tiefere Töne auftauchen, bei der Flöte höhere. In der Psychophysik hat man wegen dieser Bedeutung den Begriff des „Adaptationsniveau“ geprägt. Das Adaptationsniveau ist also das Maß für die Lage eines Bezugssystems. Eine entsprechende Theorie wurde von Harry Helson entwickelt.

Was ist das Adaptationsniveau in unserem Versuch? Es ist der Reiz, der zu dem Urteil „mittel“ führte, also zu einer „5“ auf unserer Skala. In Abbildung 33 lässt sich ablesen, dass das Adaptationsniveau bei diesem Versuch etwa bei 775 Hz lag. Stimmt das mit Ihren Daten überein?

Es ist interessant, dass das Adaptationsniveau bei unserem Versuch nicht bei dem mittleren Reiz der Reizserie liegt, die wir präsentiert hatten. Wenn sich das Bezugssystem der Probanden durch die Präsentation der 15 Töne vor dem eigentlichen Versuch gebildet hätte, hätte das Adaptationsniveau eigentlich beim mittleren Ton, also bei 800 Hz liegen müssen. Die Abweichung, die wir hier finden, weist – wie oben bereits erwähnt – darauf hin, dass wohl auch das längerfristig erlernte Bezugssystem der Probanden für Tonhöhe bei der Beurteilung der Töne eine Rolle spielte.

7.7. Diskussion: Was ist der Reiz?

Das Kapitel hatte die Überschrift „Was ist der Reiz?“. In dem Versuch waren Sie gebeten worden, die Tonhöhe verschiedener Reize als Absoluturteil anzugeben. Bei der Auswertung und Interpretation war deutlich geworden, dass es nur bedingt zielführend ist, die physikalischen

Eigenschaften von Reizen anzugeben. Für Absoluturteile ist es vielmehr viel wichtiger, das Bezugssystem der Probanden zu kennen. Ein Absoluturteil entsteht, indem eine bestimmte erlebte Ausprägung (z.B. der Tonhöhe) in Beziehung gesetzt wird zu dem Wissen darum, welche Ausprägungen von Tönen in einem bestimmten Kontext (im Versuch mit 5 unterschiedlichen Tönen, oder bei einem bestimmten Musikinstrument) auftreten können. Damit steht nicht mehr die Beziehung zwischen physikalisch gemessenen Ausprägungen bestimmter Eigenschaften und der Ausprägung des Erlebens im Vordergrund. Vielmehr geht es darum, wie eine bestimmte erlebte Reizgröße in vorhandenes Wissen (das Bezugssystem) eingeordnet wird. Als ein wichtiges Maß zur Beschreibung des Bezugssystems hatten wir das Adaptationsniveau kennengelernt, das die Lage des Bezugssystems beschreibt.

Die Frage nach dem Reiz ist damit von einer psychophysischen zu einer ganz psychologischen Frage geworden. Das Erleben eines Reizes hängt entscheidend von dem Vorwissen, der Erfahrung und dem Kontext ab. Da wir überhaupt keine Probleme damit haben, verschiedene Eigenschaften von Objekten mit Absoluturteilen zu beschreiben, müssen wir über eine Vielzahl von Bezugssystemen verfügen, die die Grundlage für diese Urteile darstellen. Die Entstehung und Nutzung dieser Bezugssysteme bekommen damit eine besondere Bedeutung, der wir in den nächsten Kapiteln weiter nachgehen werden.

8. Versuch Adaptationsniveau IIa und IIb: Der Einfluss von Anker

8.1. Anker als eine besondere Art von Kontext

Im letzten Kapitel hatten wir gesehen, dass Absoluturteile nur im Rahmen von Bezugssystemen Sinn machen. Ein Bezugssystem ist durch die Pole geprägt. Die Pole geben den Bereich an, in dem die Reizausprägungen in einem bestimmten Kontext typischerweise liegen. Die Lage des Bezugssystems wird durch das Adaptationsniveau beschrieben. Dieses ist die Mitte zwischen den Polen. Auf diese Weise passt sich die menschliche Wahrnehmung optimal an die möglichen Reizausprägungen an. Wenn man weiß, dass es um Flötentöne geht, kann man sich auch vorstellen, was ein sehr sehr tiefer Ton in diesem Kontext bedeutet. Dieser ist natürlich physikalisch ganz verschieden von dem sehr sehr tiefen Ton eines Basses. Diese Anpassung der Wahrnehmung an den Kontext ist auch sehr sinnvoll. Auf diese Weise nutzen wir die begrenzte Unterscheidungsfähigkeit unserer Sinne optimal aus, indem wir uns jeweils an den Kontext anpassen oder adaptieren.

Dieses Bezugssystem muss nicht nur den Bereich möglicher Ausprägungen umfassen, sondern kann auch komplexere Eigenschaften berücksichtigen. Bei der Schätzung der Münzgrößen im ersten Kapitel hatte sich gezeigt, dass eben nicht nur die Größe, sondern auch der Wert eine Rolle spielt. Der geringe Wert der 1-Cent-Münze hatte dazu geführt, dass deren Größe unterschätzt wurde. Das Bezugssystem für Münzgrößen enthält demnach nicht nur die Größe, sondern eben auch den Wert. Damit stellt sich jeweils die Frage, was eigentlich den Kontext ausmacht, oder was das Bezugssystem beeinflusst.

Im letzten Kapitel hatten wir ganz grundlegend untersucht, welche Rolle die Reizserie spielt, also künstlich einen Ausschnitt von Reizen hergestellt und für diesen untersucht, wie verschiedene Reize in dieses System eingeordnet werden. Dies war nur begrenzt gelungen, weil das sehr gut gelernte Bezugssystem für Tonhöhen störend gewirkt hatte. Wenn wir die Serie deutlich häufiger wiederholt hätten, hätten Sie dieses Bezugssystem immer besser erlernt und sich in Ihren Urteilen noch stärker an dieses System anpassen können – aber das wäre vermutlich von der Versuchsdauer her für Sie wenig akzeptabel gewesen.

Was macht den Kontext, der das Bezugssystem bestimmt, also aus? Der wesentliche Kern ist das Vorwissen. Man hat bestimmte Erwartungen darüber, welche Ausprägungen der Tonhöhe in der aktuellen Situation vorkommen können. Diese Erwartungen können an die tonerzeugenden Objekte gebunden sein. Man weiß, welche Töne von einem Klavier, einer Flöte oder einem Bass produziert werden. Sie können auch an die aktuelle Situation gebunden sein: Man hat Ihnen die vorkommenden Töne im Versuch vorgestellt. Wenn Sie Musik in einem bestimmten Stil hören (z.B. HipHop), wissen Sie auch ungefähr, welche Töne vorkommen werden. Und natürlich verändern sich Ihre Erwartungen in einer bestimmten Situation, wenn plötzlich Tonhöhen vorkommen, die Sie eigentlich nicht erwartet hatten.

Diese letzte Überlegung hatten wir in dem zweiten Versuchsteil verfolgt. Sie hatten wieder die Reizserie zunächst kennengelernt, sodass bei Ihnen bestimmte Erwartungen entstanden sind. Dann wurde eine Veränderung im Kontext eingeführt. Als Achtung-Signal wurde entweder ein deutlich höherer (990 Hz) oder deutlich tieferer Ton (630 Hz) eingeführt. Beim Versuch Adaptationsniveau I war der Achtung-Ton aus der Mitte der Serie gewählt worden (800 Hz). Sie waren instruiert worden, diesen Ton nicht zu beurteilen, sondern sich ganz auf den zweiten Ton zu konzentrieren. Dennoch stellt dieser Anker-Ton einen Reiz dar, der in den beiden Versuchen Adaptationsniveau II in dem Erlebensbereich „Tonhöhe“ eine bestimmte, extreme Ausprägung aufweist, entweder eine sehr tiefe (Adaptationsniveau IIa) oder eine sehr hohe (Adaptationsniveau IIb). Da diese in dem Versuch unmittelbar präsent war, war zu vermuten, dass ein solcher Anker Ihre Erwartungen über die in diesem Kontext vorkommenden Tonhöhen verändert. Dies müsste dann auch Ihr Erleben verändern.

Dieses Kapitel beschäftigt sich also nicht mit der Schifffahrt. „Anker“ sind in der Psychologie bestimmte Reize, die an den Grenzen oder außerhalb eines Bezugssystems liegen, aber die Urteile verändern können, d.h. Auswirkungen auf das Bezugssystem haben. Welche Auswirkungen haben Sie? Um diese Frage zu beantworten, werten wir im Folgenden den zweiten Versuchsteil aus.

8.2. Ergebnisse des Versuchs Adaptationsniveau II

Das Ergebnis für eine Gruppe von Studenten ist in Abbildung 34 dargestellt. Die Auswertung erfolgte in derselben Weise wie im ersten Versuchsteil. Sie haben Ihre Urteile aus den drei Versuchen in drei Tabellenblätter in Excel eingetragen. Jetzt können Sie für die beiden Versuche mit Anker ebenfalls die Mittelwerte berechnen. Tragen Sie diese in das vierte Tabellenblatt ein und erstellen Sie dann mit Ihren Daten eine Abbildung wie Abbildung 34. Dort sind die Mittelwerte in Abhängigkeit von der Tonhöhe in Hz dargestellt. Fügen Sie auch eine entsprechende Beschriftung (Legende) in Ihre Abbildung ein.

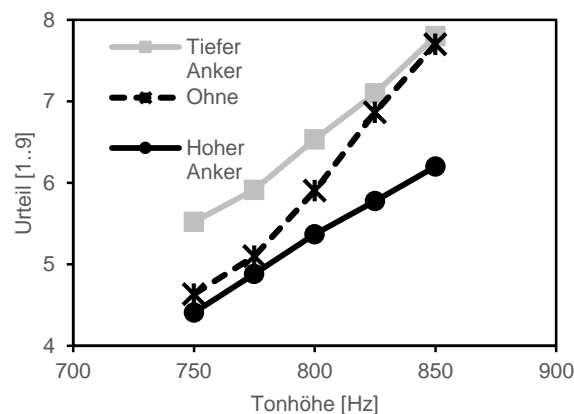


Abbildung 34: Ergebnisse des Versuchs für eine Gruppe von Studenten ($n=23$) für den ersten Versuchsteil ohne Anker und die beiden Versuchsteile mit tiefem bzw. hohem Anker.

In den Ergebnissen der Gruppe zeigte sich ein sehr deutlicher Einfluss beider Anker. Die Anwesenheit des tiefen Ankers (obere graue Linie) führt dazu, dass vor allem die tiefen Töne höher erlebt werden. Die graue Linie liegt z.B. bei 750 Hz deutlich höher als die Linie des Versuchs ohne Anker. Bei 750 Hz war das Urteil ohne Anker 4.5 („mittel tief“ bis „tief“). Mit tiefem Anker ist eine 5.5 zu finden, also „mittel“ bis „mittel hoch“. Bei den beiden höchsten Tönen wirkte sich der tiefe Anker dagegen relativ wenig aus. Ein umgekehrter Effekt ist für den hohen Anker zu finden (die schwarze untere durchgezogene Linie). Dieser beeinflusst vor allem das Erleben der hohen Töne. Während der höchste Ton ohne Anker als „sehr hoch“ (8) bewertet wird, wird er mit hohem Anker als „mittel hoch“ (6) bewertet. Auch hier wird der Einfluss des Ankers immer kleiner, je tiefer die Töne sind. Bei den beiden tiefsten Tönen ist kaum ein Unterschied zu finden. Zusammenfassend heißt das:

- Anker verschieben das Bezugssystem.
- Reize in der Nähe des Ankers werden stärker in Richtung „mittel“ erlebt. Bei einem tiefen Anker wirken die tiefen Töne nicht mehr ganz so tief, bei einem hohen Anker die hohen nicht mehr ganz so hoch.
- Der Einfluss des Ankers wird umso schwächer, je näher die Reize dem Pol liegen, der entgegengesetzt zum Anker liegt.

Ein solcher Effekt wird als Kontrasteffekt bezeichnet. Die Reize, die in der Nähe des Ankers liegen, werden durch den Kontrast zu dem Anker als nicht so extrem erlebt, so dass sich das Urteil entsprechend verschiebt. Je geringer die Ähnlichkeit zwischen Reiz und Anker ist, umso weniger entsteht dieser Kontrasteffekt und umso weniger beeinflusst der Anker das Erleben.

Was bedeutet dieser Kontrasteffekt für das Bezugssystem? Obwohl Sie die Instruktion hatten, den Achtung-Ton zu ignorieren, hat dieser dennoch Ihre Wahrnehmung beeinflusst. Ganz unwillkürlich ist dieser Ton doch Teil Ihres Bezugssystems geworden. Da der tiefe Anker tiefer war als der tiefste Ton des ursprünglichen Bezugssystems, hat er vermutlich den tiefen Pol ersetzt. Wenn der tiefe Ankerton jetzt der untere Pol Ihres Bezugssystems ist, werden die tiefsten Töne der Reizserie etwas höher beurteilt. Gleiches gilt für den hohen Anker. Wenn dieser präsentiert wurde, war er höher als der hohe Pol und hat deshalb diesen ersetzt. Sie haben eben die Erfahrung gemacht, dass der ursprüngliche Pol gar nicht der höchste Ton der präsentierten Reize in diesem Kontext ist, sondern dass es in diesem Kontext auch noch höhere Töne gibt.

- Der Kontrasteffekt entsteht, da die Anker die ursprünglichen Pole des Bezugssystems ersetzen.
- Das Bezugssystem wird durch Anker, die außerhalb des Bezugssystems liegen, erweitert.

Wieso und unter welchen Bedingungen geschieht dies? Wir hatten oben beschrieben, dass sich Bezugssysteme normalerweise als mögliche Ausprägungen von Eigenschaften von Objekten entwickeln. Sie wissen, welche Töne auf einer Flöte möglich sind. Interessant wäre es, ob sich das Bezugssystem für Flötentöne auch erweitern würde, wenn vor jedem Flötenton, den Sie beurteilen sollen, ein Basston gespielt wird. Wenn das Bezugssystem mit Objekten verknüpft ist (eines für die Flöte, ein anderes für den Bass), sollte unter diesen Bedingungen kein Kontrasteffekt entstehen. Sie können diesen Versuch relativ einfach selbst durchführen, wenn Sie Lust haben und neugierig sind.

Sie sehen hier auch wieder die psychologische Methodik. Da wir den Gegenstand der Psychologie, das Erleben, nicht direkt messen können, müssen wir uns Theorien überlegen, wie das Erleben funktionieren könnte. Wir hatten überlegt, dass ein Bezugssystem das Wissen darüber speichert, welche Ausprägungen bestimmte Eigenschaften von Objekten haben können. Wir wissen also zum Beispiel, welche Tonhöhen bei einer Flöte vorkommen können. Wenn wir Urteile über Flötentöne treffen, sollte dies nicht dadurch beeinflusst werden, dass wir gleichzeitig noch einen Bass hören. Diese Überlegung kann man nun experimentell prüfen und so untersuchen, ob unsere Überlegungen über das Erleben des Menschen durch die Daten bestätigt werden.

Vor dem Hintergrund dieser Überlegungen hatte sich ein Kontrasteffekt bei unserem Versuch dadurch ergeben, dass die Ankertöne ebensolche Sinustöne waren wie die Töne im Versuch. Wenn ein solch ähnlicher Ton in demselben Kontext präsentiert wird wie die Versuchstöne, gehört dieser Anker eben in diesen Kontext und erweitert das Bezugssystem. Der Versuch zeigt damit noch eine andere Eigenschaft von Bezugssystemen:

- Bezugssysteme sind kurzfristig durch den Kontext veränderbar und anpassbar.

Unsere Wahrnehmung kann das Bezugssystem so sehr schnell und flexibel an die Umgebung anpassen. Innerhalb dieses Kontexts sind dann absolute Urteile möglich. Umgekehrt sind die absoluten Urteile von Menschen aber eigentlich relative Urteile in zweierlei Sinn:

- (1) Absolute Urteile sind auf ein Bezugssystem bezogen und damit relativ.
- (2) Das Bezugssystem passt sich an den Kontext an, so dass absolute Urteile je nach Kontext variieren können.

In dieser Sichtweise können wir auch den Effekt des Münzwerts beim Größenschätzen interpretieren: Die Größe war in den Kontext „Münze mit einem bestimmten Wert“ gebracht worden. Durch diese Kontextänderung war ein anderes Bezugssystem aktiviert worden, sodass eben nicht nur die Größe beurteilt wurde.

Nebenbei: Warum mussten Sie den zweiten Versuchsteil zusammen mit dem ersten durchführen? Bei der Auswertung des ersten Versuchsteils hatten Sie gesehen, dass es eigentlich nur 5 unterschiedliche Töne gibt. Wenn Sie das vorher gewusst hätten, hätte dieses Wissen vielleicht Ihr Urteil beeinflusst. Neun Kategorien für fünf Töne – eigentlich darf man ja nicht alle Kategorien nutzen, weil es gar nicht so viele unterschiedliche Töne gibt. Diese Beeinflussung wollten wir vermeiden. Wir hatten Sie also ein wenig getäuscht bzw. Ihnen nicht alle Informationen über den Versuch vorher gegeben. Wir hatten gute Gründe dafür und auch keine Bedenken, dass Ihnen dieses Nicht-Wissen irgendwie schaden könnte. Dennoch: Wenn man Probanden täuscht, sollte man gut darüber nachdenken und es gut begründen können, warum dies notwendig ist.

Doch zurück zur Auswertung. Im letzten Kapitel hatten wir das Adaptationsniveau kennengelernt, das die Lage des Bezugssystems beschreibt. Gemessen hatten wir es, indem wir den Reiz gesucht haben, bei dem die Probanden mit „mittel“ urteilen. Im ersten Versuch lag dies ungefähr bei dem Reiz mit 775 Hz. Mit dem hohen Anker verschiebt sich das Niveau, allerdings nur ein kleines bisschen auf ungefähr 790 Hz. Dies ist dadurch zu erklären, dass der hohe Anker das Bezugssystem nach oben erweitert, so dass die Mitte des Bezugssystems eben höher liegt.

Sehr viel deutlicher ist hier der Effekt des tiefen Ankers. Mit diesem Anker ist in der Reizserie kein „mittlerer“ Reiz mehr vorhanden, sondern dieser liegt unterhalb von 750 Hz. Auch dies passt zu der Integration dieses Ankers in das Bezugssystem. Es wird auf tiefere Töne erweitert, sodass die Mitte eben auch tiefer liegt.

Damit kann man mit der Bestimmung des Adaptationsniveaus sehr gut den Effekt von Ankern beschreiben. Wenn man aber nicht alle Reizausprägungen präsentiert, die in dem bisher erlernten Bezugssystem vorhanden sind, findet man allerdings für die aktuelle Reizserie möglicherweise kein „mittel“-Urteil und kann so das Adaptationsniveau nur abschätzen (wenn man z.B. linear interpoliert). Man sieht dadurch sehr schön, dass Anker das aktuelle Bezugssystem deutlich über die präsentierte Reizserie hinaus verschieben können,

Auch dieser Versuch hat wieder gezeigt, dass sich unser Erleben von den physikalisch gemessenen Reizen unterscheidet. Je nach Kontext, mit dem hohen oder tiefen Anker, wurde derselbe Ton unterschiedlich erlebt. Oder wurde er nur unterschiedlich beurteilt? Ist das einfach ein Effekt der Sprache, dass wir unser Erleben sprachlich beschreiben müssen, es aber „eigentlich“ vielleicht doch identisch war? Das führt dann zur Frage, was eigentlich „das Erleben“ ist. Wenn das Erleben der beiden physikalisch gleichen Töne zu einem unterschiedlichen Verhalten führt, spräche das dafür, dass auch das Erleben unterschiedlich war. Sprache ist natürlich eine bestimmte Art des Verhaltens. So gesehen muss man aus den unterschiedlichen Urteilen folgern, dass auch das Erleben unterschiedlich war. Man könnte aber auch versuchen, dies mit Hilfe eines anderen Verhaltens zu prüfen. Man könnte z.B. Probanden trainieren, bei einem tiefen Ton die linke Maustaste zu drücken und bei einem hohen Ton die rechte (Wahlreaktionsaufgabe). Gibt man dann nach dem Training den hohen Ton mit einem sehr hohen Achtung-Signal (hoher Anker), müsste der Proband mit der linken Maustaste reagieren, wenn sich tatsächlich das Erleben so verschiebt, dass durch den Anker der Ton als „tief“ erlebt wird. Auch dies ist ein Versuch, den Sie relativ einfach selbst durchführen könnten. Ohne vorherzusagen, was dabei herauskommt: Es spricht insgesamt einiges dafür, dass sich durch den Kontext tatsächlich Ihr Erleben verändert.

8.3. Der Reiz – psychologisch gesehen

Im ersten Teil des Buchs hatten wir uns mit der Metrik des Psychischen beschäftigt. Es gibt ganz unterschiedliche Erlebensbereiche (Qualitäten), für die man jeweils die Unterscheidungsfähigkeit des Erlebens bestimmen kann (Quantität). Die Unterscheidungsfähigkeit hatte sich hier als eine zentrale Eigenschaft herausgestellt, die wir mit Unterschiedsschwelle und Unsicherheitsintervall beschrieben hatten. Die klassische Psychophysik untersucht dann weiter, wie der physikalisch definierte Reiz im Erleben abgebildet wird, also welcher funktionale Zusammenhang zwischen physikalischer Messung und Erleben besteht. Für diese Untersuchungen hatten wir die physikalischen Eigenschaften als UV variiert und das Erleben als AV (z.B. Beurteilung als „sehr sehr hoch“) gemessen.

Bei diesen Untersuchungen haben wir jetzt festgestellt, dass ein solcher Zusammenhang nur dann gilt, wenn der Kontext definiert wird und stabil bleibt. „Kontext“ hatten wir etwas genauer bestimmt als Bezugssystem, d.h. ein Wissen darüber, welche Reizausprägungen unter bestimmten Umständen typisch sind. Dieses Bezugssystem kann

- sich kurzfristig neu ausbilden (wenn man z.B. in einem Versuch eine Reizserie vorgibt),
- durch lebenslange Erfahrung stabilisiert sein (und dann z.B. kurzfristig neu ausgebildete Bezugssysteme stören) und
- durch Anker verschoben werden (Kontrasteffekte).

Diese Untersuchung des Kontexts hatte auch gezeigt, dass die Beschreibung von Bezugssystemen auch möglich ist, ohne die Beziehung zu physikalisch messbaren Reizen herzustellen. Wenn Menschen Absoluturteile treffen, ordnen sie Reize innerhalb ihres Bezugssystems ein. Die Position innerhalb des Bezugssystems ist für sehr viele Fragestellungen wesentlich interessanter als die Angabe der entsprechenden physikalischen Messgröße. Sind Sie als Autofahrer bei einer Nachtfahrt schon „sehr sehr müde“ und sollten lieber anhalten, weil Ihnen gleich die Augen zufallen werden? Als geübter Autofahrer und Mensch haben Sie ein sehr gut ausgeprägtes Bezugssystem für Ihre Müdigkeit. Wenn Sie selbst sich in diesem Bezugssystem als „sehr sehr müde“ einordnen, sollten Sie das auch ernst nehmen und eine kurze Pause machen.

Die oben gestellte Frage nach den Dimensionen des Erlebens ist damit letztlich eine Frage von Bezugssystemen. Sobald wir mühelos Dinge hinsichtlich bestimmter Eigenschaften beurteilen können, ist dies ein Zeichen dafür, dass wir über ein gut ausgeprägtes Bezugssystem verfügen. Ein Mensch ist groß oder klein, nett oder unfreundlich, attraktiv oder hässlich – die Liste ließe sich lange fortsetzen. Absoluturteile von Menschen sind damit eine hervorragende Möglichkeit, auf die Qualitäten und Quantitäten des Psychischen zuzugreifen.

Wie die beiden Versuche zum Adaptationsniveau zeigen, sind die Urteile, die man so erhält, immer abhängig vom Kontext, also vom Bezugssystem selbst (an welche Art von Tönen denkt man gerade, Bass oder Flöte?) und von aktuellen weiteren Reizen, die ähnliche Dimensionen aufweisen (Anker). Beurteilen Sie einmal Ihre eigene Müdigkeit, wenn ein schlafender Beifahrer neben Ihnen sitzt oder wenn dieser hellwach und munter ist! Die psychologisch-methodische Frage verändert sich vor dem Hintergrund dieser Überlegungen. Man fragt nicht mehr:

- Welche physikalischen Reizausprägungen führen zu bestimmten Urteilen (klassische Psychophysik)?

Vielmehr wird wichtig:

- Welches Bezugssystem verwendet der Proband, wenn er ein bestimmtes Urteil abgibt (Bezugssystemtheorie)?

Diese Veränderung der Fragestellung hat ganz praktische Konsequenzen, wenn man im Bereich der angewandten Psychologie die Ausprägung des Erlebens feststellen will. Wie stressig ist für einen Fahrer eine bestimmte Vorfahrtssituation an der Kreuzung? Ist das für den älteren Fahrer anstrengender als für junge Fahrer? Man kann natürlich sehr einfach mit Hilfe von Absoluturteilen danach fragen, wie anstrengend eine solche Situation war. Aber hatten die jungen und alten Fahrer dasselbe Bezugssystem im Kopf? Wie kann ich sie dazu bringen, dasselbe Bezugssystem zu verwenden?

Ohne diese Frage vollständig zu beantworten, geben unsere Untersuchungen im Rahmen dieses Buches bereits eine ganz wichtige Antwort: Man sollte Probanden die Pole des Bezugssystems präsentieren, innerhalb dessen sie antworten sollen. „Denken Sie an eine Fahrt auf einer leeren Autobahn mit 100 km/h. Es ist hell, sie haben gute Sicht, es ist kaum Verkehr. Das ist sicherlich eine sehr wenig anstrengende Situation. Und jetzt stellen Sie sich bitte vor, wie Sie müde von der Arbeit am frühen Abend nach Hause fahren. Es ist dunkel, es fängt an, zu regnen, Sie können kaum mehr was sehen. Gleichzeitig kommt jetzt noch eine ganz blöde Kreuzung, sehr viele andere Autofahrer, Fußgänger, Radfahrer, und jetzt auch noch eine Straßenbahn. Wer hat hier eigentlich Vorfahrt? Wo muss ich langfahren? Das ist eine sehr anstrengende Situation.“ Wenn die Probanden diese Pole im Kopf behalten, kann man davon ausgehen, dass sie ihre Absoluturteile in einem vergleichbaren Bezugssystem treffen.

Nur wenn die Probanden das tun, hat man überhaupt die Chance, Unterschiede zwischen relevanten Gruppen zu finden. Nur wenn alte und junge Fahrer dasselbe Bezugssystem im Kopf haben, kann man zeigen, dass dieselbe Situation für die jungen unerfahrenen Fahrer anstrengender ist als für die älteren erfahrenen Fahrer. Da junge Fahrer andere Fahrten durchführen als ältere Fahrer, kennen sie ganz andere Situationen als die älteren Fahrer. Junge Fahrer legen z.B. häufiger längere Wege auf Autobahnen und Landstraßen zurück, wo sie sehr viele eher langweilige und einfache Situationen erleben. Ältere Fahrer sind dagegen häufig in der Stadt unterwegs und lernen hier ständig schwierige Kreuzungssituationen kennen. Damit haben diese beiden Gruppen vermutlich ein ganz unterschiedliches Bezugssystem. Bei den jüngeren, unerfahrenen Fahrern ist dies durch einfache Situationen geprägt, bei den älteren durch schwierige. Dieselbe Kreuzungssituation könnte einfach deshalb von den jüngeren Fahrern als wesentlich anstrengender erlebt werden als von den älteren Fahrern, weil sie ein anderes Bezugssystem im Kopf haben. Dies kann man nur entscheiden, wenn man durch die oben dargestellte Instruktion dafür sorgt, dass die Urteile im selben Bezugssystem geschehen. Also:

- Bei Beurteilungen des Erlebens muss immer darauf geachtet werden, dass die Probanden dasselbe Bezugssystem nutzen.
- Die Vorgabe von Polen ist eine sehr gute Möglichkeit, das zu verwendende Bezugssystem zu verdeutlichen.

Da das Messen des Erlebens in vielen (angewandten) Bereichen der Psychologie von sehr großer Bedeutung ist, hat damit unsere Grundlagenuntersuchung zur Metrik des Psychischen sehr praktische Konsequenzen.

8.4. Zusammenfassung

Am Anfang des letzten Kapitels hatten wir die Frage gestellt „Was ist der Reiz?“. Wir hatten darauf hingewiesen, dass Reize nie für sich erlebt werden, sondern der Kontext immer eine Rolle spielt. In den Untersuchungen hatten wir den Kontext näher definiert und gesehen, dass Bezugssysteme als Wissen um mögliche Ausprägungen von Reizdimensionen eine zentrale Rolle spielen. Die konkrete Ausprägung einer Eigenschaft eines Reizes wird immer erlebt als Position in einem Bezugssystem von möglichen Eigenschaftsausprägungen. Dabei gibt es nicht

„das“ Bezugssystem, sondern dies wird angepasst an den jeweiligen Gegenstand (Bass vs. Flöte) und andere Reize im Kontext (Anker).

Gerade für die experimentelle Psychologie, die häufig das Erleben als abhängige Variable untersucht, ist deshalb die Kontrolle des Kontexts ein entscheidender Punkt. Wenn man zuverlässige und vergleichbare Urteile von unterschiedlichen Probanden erhalten möchte, muss man dafür sorgen, dass sie dasselbe Bezugssystem verwenden. Da andere Reize als Anker wirken und das Bezugssystem verändern können, muss man diese als Störvariablen kontrollieren und dafür sorgen, dass sie entweder nicht wirken können (Elimination von Störvariablen) oder zumindest bei allen Probanden gleich wirken (Kontrolle von Störvariablen). Im ersten Versuch zum Adaptationsniveau hatten wir deshalb als Achtung-Ton immer denselben Ton verwendet wie den, der beurteilt werden sollte (Elimination von Störvariablen). In den folgenden Versuchen hatten alle Probanden denselben Anker erhalten (Kontrolle von Störvariablen). Dieser hatte zwar Auswirkungen, aber bei allen Probanden in derselben Weise. Unterschiede zwischen den Probanden können also nicht daran liegen, dass sie unterschiedliche Bezugssysteme bei ihren Urteilen im Kopf hatten.

Und schließlich: Nach wie vor bleibt auch der klassische psychophysische Ansatz wichtig. In vielen Anwendungsbereichen ist es wichtig, wie bestimmte physikalisch gemessene Ausprägungen von Reizen erlebt werden. Ab welcher Lautstärke wird Lärm stressig? Wie stark darf das Licht gedimmt werden, so dass entspanntes Lesen noch möglich ist? Wie laut muss ein Warnton sein, damit der Fahrer schnell reagiert? Bei einer Vielzahl von Fragestellungen können die unterschiedlichen Methoden verwendet werden, die Sie im ersten Teil dieses Buchs kennengelernt haben. Da Bezugssysteme eine solche Bedeutung für die Untersuchung des Erlebens haben, werden wir jetzt in den folgenden beiden Kapiteln die Entstehung und Struktur von Bezugssystemen näher betrachten.

9. Grauersuch: Wie entstehen Bezugssysteme?

9.1. Wie viele Kategorien enthalten Bezugssysteme?

Ein Gegenstand erhält für uns sowohl Qualität als auch Quantität, indem wir ihn in ein Bezugssystem einordnen. Wenn wir ein Bezugssystem für die Größe von Menschen entwickelt haben, dann können wir bei jedem Menschen ohne Nachzudenken seine Größe beschreiben – „das ist aber ein großer Mann“, „die ist aber klein“, „der Bankräuber war mittel groß“. Diese absoluten Urteile sind möglich, weil wir ein Bezugssystem im Kopf haben. Das Bezugssystem selbst wird von uns aber gar nicht bewusst erlebt. So entsteht für uns der Eindruck, wir könnten absolut urteilen, obwohl jedes Urteil eigentlich ein Einordnen der Eigenschaftsausprägung in das Bezugssystem darstellt. Eigentlich handelt es sich also um einen Vergleich mit den möglichen Ausprägungen in unserem Bezugssystem. Aber weil dieser Vergleich nicht durch bewusstes Nachdenken entsteht und das Bezugssystem eben nicht bewusst repräsentiert ist, erscheint uns das Urteil als absolut.

Diese Überlegungen zeigen auch wieder das zentrale Problem der Psychologie: Viele psychische Prozesse können wir nicht direkt untersuchen, weil sie dem Menschen gar nicht bewusst sind. „Was sind die Pole Ihres Bezugssystems für die Tonhöhe einer Flöte?“ „Wie viele Kategorien hat Ihr Bezugssystem für die Größe von Münzen?“ Diese Fragen können Sie nicht beantworten, weil es sich um unbewusste Prozesse und Eigenschaften handelt. Mit diesem Konzept des Unbewussten ist Sigmund Freud bekannt geworden. Er hat dies aber in einem ganz anderen Sinn verwendet, als unbewusste Triebe und Motive, die unser Handeln bestimmen. Wir werden dies in einem allgemeineren Sinn: Viele psychische Prozesse, die unserer Wahrnehmung, dem Verständnis der Umwelt und auch unseren Entscheidungen zugrunde liegen, sind nicht bewusst und auch nicht bewusstseinsfähig. Egal, wie sehr wir über unser Bezugssystem nachdenken, wir können die oben gestellten Fragen nicht beantworten. Wir sind darauf angewiesen, indirekt, über beobachtbare oder bewusste Folgen dieser Prozesse auf diese zu schließen.

Das Experiment ist die wichtigste Methode, um unsere Überlegungen zu prüfen: Wenn wir zum Beispiel für unterschiedliche Instrumente tatsächlich unterschiedliche Bezugssysteme aufgebaut haben, dann müsste ein Anker, der von einem Bass gespielt wird, die Beurteilung von Flötentönen nicht beeinflussen. Diese Überlegung können wir prüfen, indem eine Gruppe von Probanden den Bass als Anker bekommt, eine zweite einen Flötenton. In der zweiten Gruppe müsste ein Kontrasteffekt auftreten, in der ersten dagegen nicht. Wenn wir dieses Ergebnis finden, spricht das für unsere Überlegung. Wir können so die Eigenschaften zwar nicht direkt messen, aber indirekt erschließen.

Doch zurück zum Bezugssystem selbst. Bisher hatten wir beschrieben, dass uns das Bezugssystem ermöglicht, Eigenschaften von Objekten absolut zu beurteilen. Ein Bezugssystem leistet aber noch mehr. Ein Gegenstand wird für uns nur zum Gegenstand, wenn wir seine Eigenschaften beschreiben könnten. Dies betrifft sowohl die Qualität („er hat eine bestimmte Haarfarbe“) als auch die Quantität („seine Haare sind blond“). Diese Zuweisung von Eigenschaften ist wiederum nur möglich, wenn wir ein Bezugssystem für diese bestimmte Qualität erworben haben. Ein Kind lernt die Bedeutung der Objekte um sich herum kennen, indem es ihre Eigenschaften und deren Ausprägungen kennenlernt. Oder psychophysisch beschrieben: indem es ein Bezugssystem entwickelt für bestimmte Eigenschaften. Bezugssysteme sind damit ganz grundlegend für das menschliche Verständnis der Welt. Bezugssysteme ermöglichen es erst, die Objekte zu kennen und in ihren Eigenschaften einzuordnen. Durch Bezugssysteme werden die Objekte erst bestimmte Objekte.

Damit dies funktioniert, müssen Bezugssysteme ein zentraler Bestandteil unseres Langzeitgedächtnisses sein. Im Langzeitgedächtnis wird unser erworbenes Wissen über Objekte und ihre

Eigenschaften gespeichert, sodass wir Objekte erkennen und bewerten können. Wir wissen jetzt, dass bei dieser Speicherung die Repräsentation unterschiedlicher möglicher Eigenschaften (Qualitäten) und ihrer typischen Ausprägungen (Quantitäten) enthalten ist. Man könnte jetzt weiter darüber nachdenken, wie die Beurteilung von Eigenschaften eines konkreten Objektes dann funktioniert. Wird z.B. die aktuelle Ausprägung mit allen gespeicherten möglichen Ausprägungen verglichen? Dann müsste ein Urteil umso länger dauern, je differenzierter ein Bezugssystem ist, je mehr Kategorien also verglichen werden müssen. Auch dies ist wieder ein Beispiel dafür, wie man indirekt auf psychische Prozesse zugreift: Man überlegt, wie dies funktionieren könnte und welche Auswirkungen dies hätte. Das würde man dann experimentell prüfen.

Ein Bezugssystem enthält das Wissen über die möglichen Ausprägungen einer Eigenschaft, die wir bislang erlebt haben. Wie entsteht dieses Wissen und wie ist dieses Wissen in unserem Gedächtnis repräsentiert? Wir haben in den vorigen Kapiteln die besondere Bedeutung von Polen, aber auch von Anker, kennengelernt. Die Extremwerte geben die Bandbreite möglicher Quantitäten an. Vermutlich sind diese auch direkt im Gedächtnis als Pole repräsentiert. Dies ist in Abbildung 35 grafisch dargestellt. Wird ein Anker präsentiert, der außerhalb dieser Pole liegt, wird vermutlich dieser Anker als neuer Pol übernommen, sodass sich das Bezugssystem verschiebt.

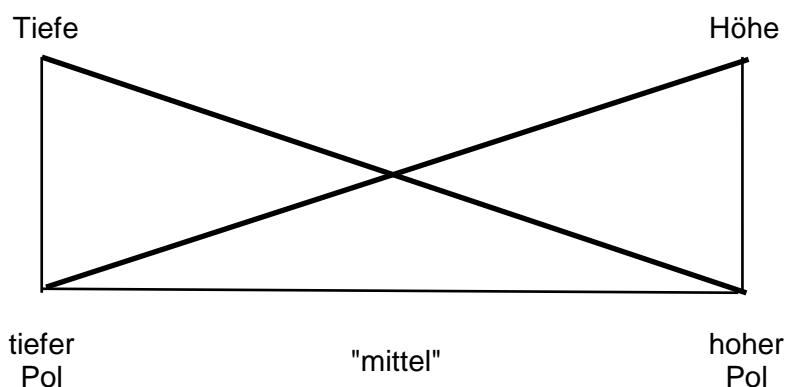


Abbildung 35: Erster Schritt bei der Entstehung eines Bezugssystems. Die beiden Pole sind repräsentiert und in der Mitte liegt zunächst ein Punkt, bei dem man sich weder für „hoch“ noch „tief“ entscheiden kann, sondern „mittel“ wählt.

Wir hatten weiter das Adaptationsniveau kennengelernt als Mitte zwischen den Polen, das die Lage des Bezugssystems repräsentiert. In Abbildung 35 wäre dies der Punkt „mittel“. Wenn man ein Bezugssystem erlernt, bildet sich möglicherweise aus diesem Punkt zwischen den Polen ein Bereich aus, der sowohl dem unteren als auch dem oberen Pol unähnlich ist, so dass dieser Bereich als dritte Kategorie im Gedächtnis gespeichert wird. Dies ist in Abbildung 36 dargestellt. Zu diesem Zeitpunkt besteht das Bezugssystem dann aus drei Kategorien, die im Gedächtnis repräsentiert werden. Dabei ist die „mittel“-Kategorie breiter als die beiden äußeren, da sie sich ja nach beiden Seiten (sowohl zu „hoch“ als auch zu „tief“) abgrenzt.

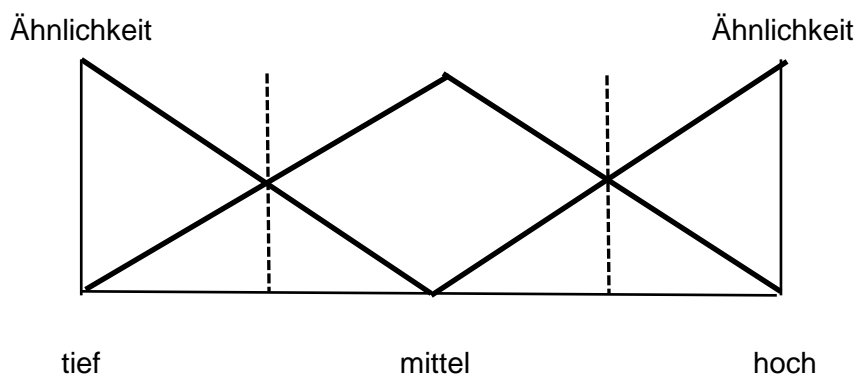


Abbildung 36: Bei der weiteren Entwicklung eines Bezugssystems entwickelt sich eine neue Kategorie „Mittel“ als Bereich zwischen den Polen „Tief“ und „Hoch“, der beiden Polen unähnlich ist.

Man kann sich nun weiter vorstellen, dass sich das Bezugssystem entsprechend diesem Prinzip mit zunehmender Erfahrung weiter ausdifferenziert. Dies ist in Abbildung 37 dargestellt. Die tiefsten Töne aus der Kategorie „tief“ werden jetzt als „sehr tief“ wahrgenommen. Die tiefsten Töne aus der Kategorie „mittel“ erscheinen jetzt als „tief“. Dasselbe passiert im Bereich der hohen Töne. Die höchsten Töne aus der bisherigen Kategorie „hoch“ erscheinen jetzt als „sehr hoch“. Die höchsten Töne aus der Kategorie „mittel“ erscheinen jetzt bereits als hoch.

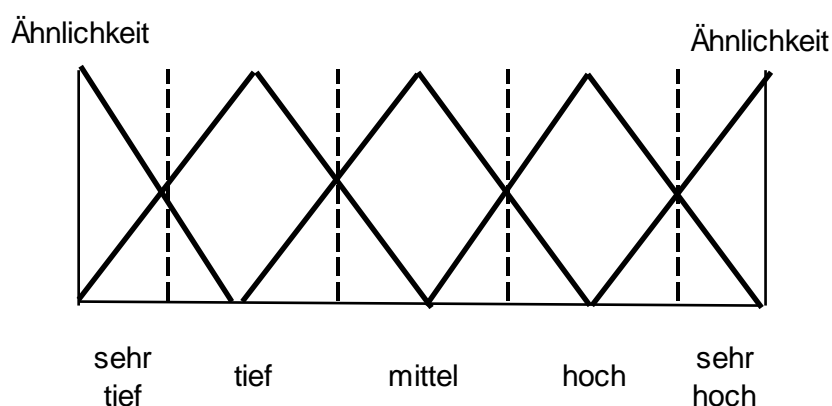


Abbildung 37: Im nächsten Schritt entwickelt sich eine Kategorie zwischen „Tief“ und „Mittel“ und zwischen „Mittel“ und „Hoch“. Zur weiteren Erklärung, s. Text.

Auch hier ist wieder zu erkennen, dass die drei mittleren Kategorien gleich breit sind, also eine ähnliche Menge von Reizen enthält. Die beiden äußeren Kategorien sind wieder etwas halb so breit, da ja nur eine Abgrenzung in eine Richtung stattfindet.

Wie lange geht diese Ausdifferenzierung weiter? Prinzipiell könnte dies so weit gehen, bis die Schwelle ebenmerklicher Unterschiede erreicht ist. Je besser die Unterscheidungsfähigkeit unserer Sinne für bestimmte Dimensionen ist, umso mehr Kategorien würde dann das Bezugssystem enthalten, die im Gedächtnis gespeichert werden müssen. Dies erscheint allerdings nicht unbedingt plausibel, da ein Bezugssystem damit im Gedächtnis sehr groß werden würde. Das würde wiederum absolute Urteile erschweren, da jeder Reiz mit allen diesen gespeicherten Ausprägungen verglichen werden müsste, um den ähnlichsten Bezugsreiz zu finden und so ein Urteil treffen so können. Wir hatten oben schon kurz beschrieben, wie man diese Überlegung experimentell prüfen könnte.

Plausibler erscheint eine begrenzte Anzahl von Kategorien, wobei eine interessante Frage ist, wie viele Kategorien typischerweise gespeichert werden. Vermutlich hängt die Anzahl von Kategorien auch von der Erfahrung ab. Es scheint eine Legende zu sein, dass Eskimos viel mehr Wörter für Schnee haben als Europäer. Modebewusste Frauen können allerdings häufig viel präziser die Farbe eines Stoffes beschreiben als an der Mode desinteressierte Männer. Aber das ist eine Frage, die hier nicht weiter vertieft werden sollen. Vielleicht möchten Sie ja eine eigene Untersuchung zu diesem Thema durchführen. Halten wir an dieser Stelle fest, dass das Bezugssystem vermutlich aus einer begrenzten Anzahl von Kategorien besteht, was eine hinreichend genaue Differenzierung des Erlebens ermöglicht.

Aber zurück zur Entstehung des Bezugssystems. Stimmen die Überlegungen zur Entstehung von Bezugssystemen durch zunehmende Differenzierung, so wie es in Abbildung 35, Abbildung 36 und Abbildung 37 dargestellt ist? Wie könnte man prüfen, ob das stimmt?

Wenn Sie Abbildung 37 betrachten, wird deutlich, dass bei einer solchen Entstehung eines Bezugssystems die drei Kategorien in der Mitte jeweils ähnlich große Bereiche von Reizausprägungen enthalten. Die beiden äußeren Kategorien sind jeweils nur ungefähr halb so groß und enthalten ungefähr die Hälfte der Reizausprägungen. Wenn man dieses Muster (gleich große Kategorien in der Mitte, halb so große Kategorien an den beiden Enden) auch in einem Versuch finden würde, dann wäre dies ein Hinweis, dass unsere Überlegungen stimmen.

Dies wollen wir im folgenden Versuch untersuchen. Wir wechseln dabei wieder die Modalität und untersuchen das Sehen. Die Qualität, die uns interessiert, ist die „Grauheit“ eines Reizes. Sie haben in Ihrem Leben sehr viele weiße und schwarze Gegenstände kennengelernt und das Wetter in Braunschweig ist häufig grau, so dass Sie sicherlich auch viele Grauschattierungen kennen. Sie sollten also über ein gut ausdifferenziertes Bezugssystem verfügen. Sind in diesem die mittleren Kategorien etwa gleich groß und bei beiden äußeren Kategorien etwa halb so groß?

9.2. Instruktion Grauersuch

In diesem Versuch sollen Sie die Grauheit eines Quadrats auf dem Bildschirm beurteilen. Dabei steht Ihnen folgende Skala zur Verfügung:

Schwarz
Sehr dunkles Grau
Dunkles Grau
Mittel Grau
Helles Grau
Sehr helles Grau
Weiß

Wie Sie sehen, sind die beiden Pole in der Skala enthalten. Wir werden auch auf dem Bildschirm immer die beiden Pole zeigen, da bei jedem Computerbildschirm Schwarz und Weiß je nach Kontrast und Helligkeit unterschiedlich aussehen können. Auf diese Weise haben Sie beim Urteilen immer die beiden Pole vor Augen. Die Reizserie umfasst diese beiden Reize (ein reines Weiß und ein reines Schwarz) allerdings nicht.

Sie werden insgesamt 115 Quadrate beurteilen. Darin sind einige Wiederholungen enthalten. Diese große Zahl von Quadraten ist notwendig, um hinterher wirklich entscheiden zu können, ob die Kategorien des Bezugssystems gleich groß sind. Wir geben aber wieder vor dem Versuch nicht genau an, wie viele unterschiedliche Graustufen Sie sehen werden, damit dieses Wissen

nicht Ihr Urteil beeinflusst. Es sind allerdings keine völlig schwarzen oder völlig weißen Quadrate dabei. Sie müssen also nicht die Pole für Ihre Urteile benutzen. Aber wenn Ihnen ein Quadrat wirklich völlig schwarz oder völlig weiß und gar nicht grau erscheint, dann dürfen Sie die Pole nutzen.

Sorgen Sie bitte für gute Beleuchtung und guten Kontrast Ihres Bildschirms. Arbeiten Sie auch hier wieder ungestört und konzentriert. Wenn Sie merken, dass Sie ermüden, machen Sie eine kurze Pause, bevor es weiter geht. Starten Sie jetzt das Programm, wählen Sie „Grauversuch“ und kopieren Sie am Ende dann Ihre Werte nach Excel.

Und hier geht es zum Versuch:

- <https://cgi-psych.rz.tu-bs.de>

9.3. Auswertung Grauversuch

Bei dem Versuch wurden 23 Reize verwendet, wobei ein reines Weiß und ein reines Schwarz nur als Pole gezeigt wurden, aber nicht in der Reizserie enthalten waren. Das Programm hat die Reize schon sortiert. In den fünf Zeilen finden Sie Ihre fünf Urteile für jeden Reiz. Bei jedem Reiz ist der Weißanteil in Prozent angegeben. Es geht von 6% bis 94%. In der Vorlage finden Sie dann weiter unten erneut die 23 Reize. Berechnen Sie für jeden den Mittelwert aus Ihren fünf Urteilen. Dabei entspricht der Wert 0 der Kategorie „schwarz“, die 6 der Kategorie „weiß“. In der Exceltabelle finden Sie eine zweite, in der die Weißanteile in Prozent in der ersten Spalte stehen, die Mittelwerte in der zweiten. Kopieren Sie einfach die Mittelwerte, die Sie berechnet haben. Dann gehen Sie in die erste Zeile in der zweiten Tabelle und fügen dort die Werte ein. Beim Einfügen wählen Sie zunächst „Einfügen - Inhalte einfügen“. Bei der Auswahl ganz wichtig: Markieren Sie links oben „Werte“ (dann werden nur die Zahlen eingefügt, aber nicht die Formeln) und „Transponieren“ (um Zeilen und Spalten zu vertauschen).

Diese Mittelwerte tragen Sie dann über den Reizen (6 – 94) auf, wie es in Abbildung 38 für 77 Studierende des WS 23/24 dargestellt ist. Ist Ihre eigene Funktion auch so schön linear? Wenn man einen einzelnen Probanden darstellt, sieht man manchmal ein bisschen fünf Stufen, was den fünf genutzten Kategorien (ohne 0 und 6) entspricht. Wenn man hier über mehrere Versuchspersonen mittelt wie in der Abbildung 38, verschwinden diese „Treppen“ weitgehend.

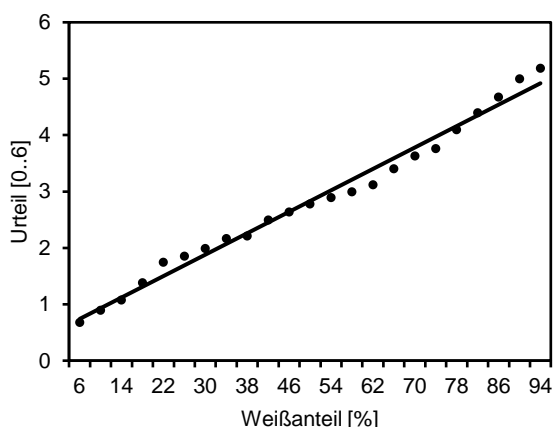


Abbildung 38: Ergebnis des Grauversuchs für $n = 77$ Studierende aus dem WS 23/24. Auf der x-Achse sind die Reize angeordnet nach dem Weißanteil (von schwarz mit 6% bis zu weiß mit 94%). Auf der y-Achse sind die Mittelwerte der Urteile dargestellt (0 = schwarz, 6 = weiß). Die Linie zeigt eine lineare Anpassung.

Wenn man die 23 Reize tatsächlich voneinander unterscheiden kann, sollten die Mittelwerte ansteigend sein. Denken Sie zurück an die verschiedenen Skalenniveaus. Das erste war das

Nominalniveau, genau diese Unterscheidbarkeit. Wenn Sie die Graustufen unterscheiden können, sollte sich auch die zunehmende Weißheit in den Urteilen zeigen. Deshalb sollten die Mittelwerte eben ansteigend sein. Dies entspricht dem oben beschriebenen Ordinalniveau. Wenn gleiche physikalische Abstände auch gleich erlebt werden, sollte der Zusammenhang annähernd linear sein. Wie die Abbildung zeigt, ist auch dies der Fall. Damit sind die Mittelwerte auf dem Intervallniveau interpretierbar.

Vielleicht fragen Sie sich, wie man eigentlich mit nur 7 (bzw. 5 ohne schwarz und weiß) Kategorien trotzdem erkennen kann, dass die 23 Graustufen unterschieden werden können. Dies ist nur deshalb möglich, weil Sie jeden Reiz mehrfach beurteilt haben. Beim hellsten Reiz waren Sie sich vermutlich sehr sicher und haben z.B. fünf Mal die Kategorie 5 gewählt. Beim nächsten Reiz waren Sie sich nicht mehr ganz so sicher, sodass Sie 4 Mal 5 und einmal 4 gewählt haben. Im Mittelwert erhält dann dieser Reiz einen etwas niedrigeren Wert. Auf diese Weise kann man also auch mit einer begrenzten Zahl von Kategorie sehr gut nachweisen, dass man viel mehr Abstufungen der Reize unterscheiden kann als Kategorien vorhanden waren. Das funktioniert noch besser, wenn Sie auch noch Mittelwerte über mehrere Versuchspersonen bilden, weil dann eben noch mehr Urteile für jeden einzelnen Reiz vorhanden sind.

Wie die eingezeichnete Linie zeigt, besteht ein annähernd linearer Zusammenhang zwischen der physikalisch abgestuften Graueit und der erlebten Graueit. Die im Versuch verwendeten gleichen Abstufungen werden also auch entsprechend gleichabständig durch die Versuchspersonen erlebt. Man sieht auch, dass die Pole teilweise gewählt wurden, weil der Mittelwert für den dunkelsten Reiz unter 1 lag, für den hellsten Reiz ein bisschen über der 5. In den meisten Fällen war aber Reiz mit 6 % noch kein Schwarz, sondern nur ein sehr dunkles grau. Der Reiz mit 94 % war ebenso noch kein weiß, sondern nur ein sehr helles grau. Die Pole gehören nicht zur Reizserie, sondern stehen außerhalb.

Mit Hilfe dieser Grafik mit den Mittelwerten können wir allerdings noch nicht über unsere Versuchsfrage entscheiden. Dazu müssen wir die Daten auf eine andere Weise aufbereiten und für jeden Reiz auszählen, wie häufig bei diesem Reiz die verschiedenen Urteilkategorien benutzt wurden. Dies wurde in Excel schon vorbereitet mit einer kleinen Formel. Gehen Sie einfach auf die Tabelle „HäufigkeitstabelleGrauversuch“. Dort finden Sie oben von Excel ausgezählt die Häufigkeiten der einzelnen Kategorien für jeden Reiz.

Um eine Grafik zu erstellen, sollten Sie diese Werte noch kopieren und transponieren. Auch dafür ist schon eine Tabelle in Excel vorbereitet. Hier stehen in den Spalten die 5 Kategorien, in den Zeilen die verschiedenen Reize. Mit dieser Tabelle können Sie dann eine Grafik wie in Abbildung 39 erstellen. Hier sind die verschiedenen Urteile (Kategorien) verschiedene Serien von Daten, die über dem Weißanteil aufgetragen werden. In der Abbildung ist das Ergebnis für die 77 Studierenden des WS 23/24 dargestellt. Die Häufigkeit wurde hier in Prozentwerte umgerechnet. Man erkennt, dass bei 6 %, 10% und 14 % sehr selten auch „schwarz“ geurteilt wurde und bei 90 % und 94 % sehr selten „weiß“. Aber betrachten wird dann das Muster der anderen Urteile.

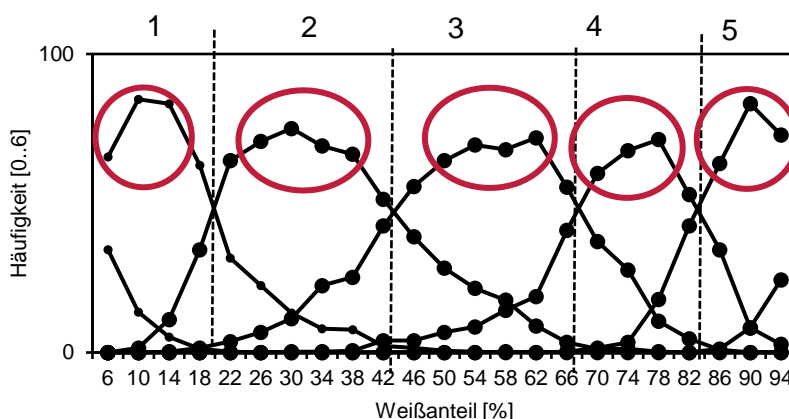


Abbildung 39: Häufigkeit der Urteile (y-Achse) pro Reiz (x-Achse). Die verschiedenen Linien stellen die verschiedenen Kategorien dar. Da in den Versuchen fünf Durchgänge durchgeführt wurden, kann die Häufigkeit maximal 5 erreichen.

Die gestrichelten Linien sind nachträglich per Hand an den Schnittstellen der Häufigkeitsfunktionen eingezeichnet und geben damit die Breite des Reizbereichs an, für den jeweils ein Urteil dominiert. Fügen Sie bei Ihren Daten entsprechende Linien ein!

Man sieht, dass diese Linien teilweise sehr nahe an einzelnen Reizen liegen, wo man sich nicht klar entscheiden kann, ob diese Reizausprägung zu der einen oder anderen Kategorie gehört. Mit den roten Kreisen wurden daher die Ausprägungen markiert, die klar in eine Kategorie gehörten.

Die Kategorie 1 „sehr dunkles Grau“ enthält die ersten drei Reize, die Kategorie 2 „dunkles Grau“ vier oder fünf Reize, die Kategorie 3 „mittleres Grau“ ebenfalls 4 oder 5 Reize, die Kategorie „helles Grau“ 3 Reize und die letzte Kategorie „sehr helles Grau“ 3 Reize. Dieses Muster entspricht nicht ganz perfekt unserer Vorhersage, aber liegt doch sehr nahe daran:

- Die mittleren Kategorien sind annähernd gleich breit, enthalten also eine ähnliche Anzahl von Reizausprägungen.
- Die beiden äußeren Kategorien sind enger und enthalten weniger Reizausprägungen.

Auch wenn dieses Ergebnis nicht perfekt mit den Überlegungen übereinstimmt, so scheint es doch sehr nahe an der Vorhersage zu sein. Die Überlegung, dass sich ein Bezugssystem so ausdifferenziert, wie es oben beschrieben wurde, scheint also plausibel zu sein. Je mehr Erfahrungen wir mit einer Reizqualität sammeln, umso feiner und differenzierter wird das Bezugssystem, wobei die einzelnen (mittleren) Kategorien annähernd ähnlich breit sind. Nur die äußeren Kategorien sind deutlich schmaler.

In dem Versuch waren ein reines Schwarz oder ein reines Weiß nicht gezeigt worden. Es ist allerdings zu vermuten, dass wenn man diese miteinschließen würde, diese Kategorien auch nur für diese Pole verwendet werden würden (zumindest wenn der hellste und dunkelste Reiz eben merklich unterschiedlich sind von einem Schwarz oder Weiß). Die Pole des Bezugssystems scheinen von daher etwas Besonderes zu sein. Sie entsprechen nicht einem breiten Bereich von Reizen, die ihnen ähnlich sind, sondern sind qualitativ etwas anderes als die anderen Abstufungen. Nur wenn andere Reize von den Polen nicht unterscheidbar sind, werden Sie in diese Kategorie eingeordnet. Bereits leichte Unterschiede führen dazu, andere Kategorien zu verwenden. Auch ein ganz ganz helles Hellgrau ist eben kein Weiß. Auch ein ganz ganz dunkles Dunkelgrau ist noch kein Schwarz. Weiß und Schwarz scheinen damit eine andere Qualität zu haben als Grau.

Damit entsteht eine paradoxe Situation: Ein Bezugssystem hilft uns, die verschiedenen Ausprägungen einer Qualität einzuordnen. Besonders wichtig sind dabei die Pole, weil sie die extremen möglicherweise auftretenden Ausprägungen enthalten und damit die Urteile begrenzen. Genau diese Pole haben aber eine eigene Qualität und gehören damit eigentlich gar nicht zum Bezugssystem! Wir finden dieses Phänomen bei sehr vielen Qualitäten und man muss sich bei jeder Qualität, die beurteilt werden soll, von Neuem fragen, ob die Pole zur Reizserie gehören oder qualitativ verschieden sind. Man sollte dem auch gerecht werden, wenn man Probanden eine Skala vorgibt. Die oben verwendete Kategorienskala hätte also eigentlich z.B. so aussehen sollen:

Schwarz

Sehr dunkles Grau
Dunkles Grau
Mittel Grau
Helles Grau
Sehr helles Grau

Weiß

Auf diese Weise wird auch grafisch ganz deutlich gemacht, dass die Pole Schwarz und Weiß etwas Besonderes sind. Wir haben oben darauf verzichtet, dies so darzustellen, da man dann hätte argumentieren können, dass die Ergebnisse nur durch diese Art der Darstellung so aussehen – wenn die Pole als so besonders dargestellt werden, dann wählt man diese Kategorien eben nur sehr selten.

Allerdings hatten wir im Versuch die beiden Pole auch immer auf dem Bildschirm gezeigt. Das könnte dazu verleiten, gerade in Bezug auf die Pole eben kein Absoluturteil zu treffen. Wenn ein Reiz auftaucht, der „schwarz“ aussieht, kann man durch den direkten Vergleich mit dem Pol daneben entscheiden, ob er wirklich schwarz ist. Man hat so ein Vergleichsurteil hinsichtlich der Pole provoziert. Möglicherweise wäre das Ergebnis anders ausgefallen, wenn die Pole nicht sichtbar gewesen wären.

Sie könnten das sehr einfach prüfen, indem Sie die entsprechenden Bildschirmbereiche abkleben und den Versuch mit Probanden erneut durchführen. Wir haben uns allerdings dafür entschieden, diese beiden Pole zu zeigen, da die Darstellung auf dem Bildschirm von der Helligkeit und dem Kontrast her begrenzt ist. Unabhängig vom jeweiligen Bildschirm haben wir durch die Darstellung der Pole dafür gesorgt, dass jeder Proband das gleiche Bezugssystem verwendete. Diese „Eichung“ schien uns wichtiger zu sein als mögliche negative Effekte.

9.4. Zusammenfassung

Wir hatten in diesem Kapitel untersucht, ob unsere Überlegungen, wie ein Bezugssystem entsteht, zutreffen. Wir hatten dazu ein Modell stetiger Ausdifferenzierung vorgeschlagen. Dabei wird als Bezugssystem eine begrenzte Menge von typischen Ausprägungen der entsprechenden Eigenschaft gespeichert. Durch die Ausdifferenzierung wird das erlebte Reizkontinuum in gleich große Bereiche aufgeteilt, für die jeweils ein Prototyp im Gedächtnis gespeichert ist. Nur die beiden äußeren Bereiche sind enger, da sie sich nur in eine Richtung abgrenzen müssen. Diese Menge der möglichen Reizausprägungen ist durch die Pole nach unten und oben begrenzt. Wenn man ein Kontinuum von Reizen vorgibt, müssten nach dieser Vorstellung die einzelnen (mittleren) Kategorien, mit denen man die Reize beurteilt, gleich häufig verwendet werden. Nur die beiden äußeren Kategorien sollten bei weniger Reizausprägungen verwendet

werden, da sie sich nur aus einer Richtung differenzieren. Dies war also das Entscheidungskriterium: Werden die mittleren Kategorien gleich häufig verwendet und die beiden äußeren Kategorien seltener?

Genau das zeigte sich im Versuch. Jede der drei mittleren Kategorien, die Graueit beschreiben, wurde praktisch gleich häufig verwendet. Die beiden äußeren Kategorien wurden seltener verwendet. Insgesamt erscheint damit dieses Modell der Entstehung von Bezugssystemen brauchbar. Mit zunehmender Erfahrung mit bestimmten Qualitäten differenziert sich das entsprechende Bezugssystem immer weiter aus. Allerdings müssen wir aufgrund der Ergebnisse mit den Polen unsere Vorstellung vom Bezugssystem etwas verändern: Die Pole scheinen in dem Bezugssystem anders repräsentiert zu sein als die anderen Kategorien zwischen den Polen. Jeder Pol steht für eine bestimmte Reizausprägung. Sobald wir erkennen können, dass ein bestimmter Reiz sich vom Pol in dieser Ausprägung unterscheidet, ist dieser Reiz kein Pol mehr.

Bei den anderen Kategorien scheint es anders zu sein: Ein Bereich von Reizen, der in sich durchaus unterscheidbar ist, erscheint wohl dieser Kategorie hinreichend ähnlich, so dass er dieser zugeordnet wird. Dies erkennt man an der linearen Abhängigkeit der Urteile von den Reizausprägungen. Gleichzeitig erscheint diese Überlegung auch plausibel: Wenn man als Mensch jede unterscheidbare Ausprägung einer bestimmten Qualität speichern würde, würde es bei jeder Beurteilung sehr lang dauern, bis man beim mentalen Vergleich genau diese Ausprägung gefunden hat und dann das Urteil abgeben kann. Es erscheint wesentlich effektiver, nur eine begrenzte Anzahl von Kategorien zu unterscheiden und für jede mental einen Prototyp zu speichern. Ein Urteil ist dann ein Vergleich mit diesen Prototypen. Je ähnlicher ein Reiz ist, desto wahrscheinlicher ist es, dass man den Reiz diesem Prototyp zuordnet. Wenn ein Reiz genau zwischen zwei Prototypen liegt, wird man sich vermutlich in der Hälfte der Fälle für den einen bzw. den anderen Prototypen entscheiden.

Man sollte dabei aber nicht vergessen, dass dieses Bezugssystem nicht statisch ist. Wie die Ankerversuche im Kapitel 8 gezeigt haben, kann durch die Anwesenheit von Anker das Bezugssystem kurzzeitig verschoben werden. Wenn diese Anker sehr häufig auftreten, können sie auch zu den neuen Polen eines dann in der Breite veränderten Bezugssystems werden. Ansonsten zeigen diese Ankereffekte, welche Rolle der Kontext bei Urteilen spielt. Wir haben also auf der einen Seite gelernt, welche Ausprägungen von Reizen vorkommen können und können diese deshalb absolut beurteilen. Dabei spielt der Kontext aber immer eine wichtige Rolle. Dies zeigt zusammenfassend Abbildung 40 erneut.

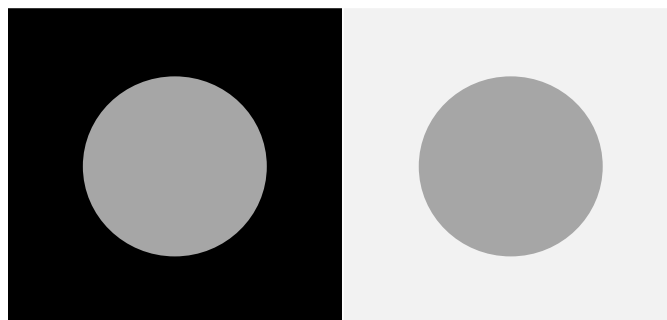


Abbildung 40: Kontextabhängigkeit der Wahrnehmung. Zur weiteren Erklärung s. Text.

Zusammenfassend lassen sich also vier Prozesse bei der Entstehung von Bezugssystemen unterscheiden:

- 1) Das Ausbilden neuer Systeme, neuer Dimensionen oder Qualitäten
- 2) Die Ausdifferenzierung dieser Systeme mit Erfahrung
- 3) Die kontextabhängige Verschiebung der Systeme

4) Die Erweiterung der Systeme (Polverschiebung) bei neuen Erfahrungen

Bei diesen Untersuchungen haben wir auch eine ganze Reihe über die Eigenschaften der Bezugssysteme gelernt:

- 1) Es gibt einen unteren und oberen Pol, der jeweils als Prototyp gespeichert ist.
- 2) Reize entsprechen nur dann dem Pol, wenn sie von ihm ununterscheidbar sind.
- 3) Zwischen den Polen ist eine begrenzte Anzahl von Erlebensbereichen (Kategorien) vermutlich ebenfalls in Form jeweils eines Prototyps pro Kategorie gespeichert.
- 4) Urteile erfolgen durch Vergleich mit den Prototypen. Je ähnlicher sie diesen sind, umso wahrscheinlicher ist es, dass sie dieser Kategorie zugehörig beurteilt werden.

Bei dieser Beurteilung scheinen zufällige Schwankungen des Erlebens eine gewisse Rolle zu spielen. Ansonsten wäre nicht zu erklären, dass ein und derselbe Reiz einmal der einen, ein anderes Mal der anderen Kategorie zugeordnet wird. Man findet bei physiologischen Untersuchungen eine mögliche Ursache dafür: Derselbe Reiz löst nicht immer dieselbe physiologische Reaktion aus, sondern die Stärke der Reaktion variiert zufällig. Ausführliche Beschreibungen dazu finden Sie in Büchern zu den physiologischen Grundlagen der Wahrnehmung.

Dieses Wissen um Bezugssysteme, die der Quantität des menschlichen Erlebens zugrunde liegen, ist auch für angewandte arbeitende Psychologen wichtig. Sie sind in vielen Bereichen darauf angewiesen, dass andere Personen ihnen ihre Erlebnisse beschreiben. Wie stark ist die Traurigkeit des depressiven Patienten? Wie groß ist der Stress am Arbeitsplatz? Wie müde ist der Autofahrer? Man erhält nur dann vernünftige Antworten auf diese Fragen, wenn es einem gelingt, die Fragen so zu stellen, dass die Probanden auf ihr vorhandenes Bezugssystem zurückgreifen können. Und die Probanden müssen mit Kategorien antworten können, die möglichst gut den mentalen Kategorien ihres Bezugssystems entsprechen. Um die Antworten einordnen und vergleichen zu können, muss man sicherstellen, dass alle Probanden dasselbe Bezugssystem verwenden. Dazu muss man den Kontext, insbesondere das Vorliegen von Ankerpunkten, kontrollieren. Bei der Entwicklung eines Fragebogens sind damit drei Aspekte besonders wichtig:

- 1) Die Formulierung der Frage: Wie fragt man so, dass der Proband auch die richtige Qualität beantwortet? Was meint man z.B. mit „Stress“? Körperliche Belastung, geistige Belastung, Zeitdruck?
- 2) Die Vorgabe von Antwortkategorien: Die Abstufungen der Qualitäten sind im Gedächtnis sicher nicht als Zahlen gespeichert, sondern vermutlich als Prototypen mit einer bestimmten Ausprägung. Verbale Kategorien, die diesen Ausprägungen entsprechend, ermöglichen es daher immer, besser auf sein Bezugssystem zuzugreifen. „Sehr müde“ ist immer eine bessere Kategorie als „6“.
- 3) Die Vorgabe von Polen: Pole bestimmen den Bereich des Bezugssystems. Wenn man direkt vor oder bei der Frage die Pole vorgibt, stellt man sicher, dass jeder Proband auf denselben Bereich des Bezugssystems zurückgreift. Die Vorgabe von Polen verbessert die Vergleichbarkeit der Urteile sehr deutlich. Auch die Pole müssen natürlich so beschrieben werden, dass sie möglichst gut dem Empfinden der Probanden entspricht.

Aber zurück zur Frage dieses Kapitels: Wir hatten unsere Vorstellung zur Entstehung der Bezugssysteme überprüft, indem wir eine Vorhersage für den Versuch abgeleitet hatten. Die Urteile sollten sich von der Häufigkeit her gleichmäßig auf die mittleren Kategorien verteilen und deutlich weniger auf die beiden äußeren Kategorien. Dies traf auch im Wesentlichen zu. Aber haben wir damit wirklich bewiesen, dass die Erlebensbereiche gleich groß sind, wie es unsere Vorstellung war? Wir konnten schließlich nicht das Erleben selbst untersuchen, sondern nur die

Urteile in den Kategorien, die wir vorgegeben hatten. Geben wir nicht das Ergebnis schon dadurch vor, dass wir diese Kategorien zeigen, die die Probanden verwenden sollen? Und wie könnte man diese Frage prüfen? Dies wird im letzten Versuch dieser Übung beantwortet.

10. Versuch Strichtafel: Wie genau sind unsere Urteile?

10.1. Bezugssystem oder Kategorien?

Im letzten Kapitel hatten wir etwas über die Entstehung von Bezugssystemen erfahren. Sie bilden sich vermutlich durch Differenzierung aus, wenn wir immer mehr mögliche Ausprägungen einer Eigenschaft kennenlernen. Als Beleg dafür hatten wir die annähernde Gleichverteilung der Urteile auf die verschiedenen Kategorien gewertet, als wir im Versuch eine Reizserie relativ feinabständig unterteilt vorgegeben haben. Aber ist das wirklich ein Beleg für die Organisation des Bezugssystems im Gedächtnis? Oder ist das nur ein Effekt unserer Skala, die wir für die Urteile vorgegeben haben? Vielleicht versucht man als Versuchsperson unwillkürlich, die Urteile auf die Kategorien gleichmäßig zu verteilen, vor allem dann, wenn die Reize auch gleichmäßig über ein Kontinuum möglicher Ausprägungen verteilt sind.

Wie kann man das prüfen? Die Alternativerklärung ist, dass Menschen in psychologischen Untersuchungen versuchen, ihre Urteile gleichmäßig auf die Kategorien zu verteilen. Man hatte die Probanden im Grauversuch ja instruiert darüber, aus welchem Bereich die Reize stammen werden, indem die Pole immer sichtbar waren. Unter diesen Bedingungen versucht man vielleicht, die vorgegebenen Kategorien möglichst gut auszunutzen. Das machte ja durchaus Sinn, da die verschiedenen Ausprägungen auch gleich häufig vorgegeben worden waren. Was würde passieren, wenn man sehr viele helle graue oder dunkle graue Reize geben würde? Wenn die Gleichverteilung ein Effekt der Kategorien ist, müssten auch unter diesen Umständen die verschiedenen Kategorien gleich häufig genutzt werden. Man würde das „Bezugssystem“ unter diesen Umständen so anpassen, dass man in dem Bereich, in dem viele Reize vorhanden sind, feiner differenziert (oder jedenfalls die Kategorien stärker ausnutzt), sodass insgesamt wieder eine Gleichverteilung auf die Kategorien entstehen würde.

Wenn die Gleichverteilung aber tatsächlich die Struktur des Bezugssystems widerspiegelt, dann müsste eine Reizserie mit vielen hellgrauen oder aber vielen dunkelgrauen Reize dazu führen, dass die entsprechenden Kategorien auch häufiger genutzt werden. Die Nutzung der Kategorien müsste also der Verteilung der Reize entsprechen. Diese Überlegung ist recht einfach zu prüfen, also auf zum letzten Experiment!

Führen Sie jetzt die Versuche Strichtafel I und II durch. Im Folgenden finden Sie die Instruktionen. Versuchen Sie bei dem Versuch, sich nicht durch die Überlegungen gerade beeinflussen zu lassen, sondern urteilen Sie möglichst spontan. Bei einem „echten“ Experiment dürfte man die Probanden natürlich nicht so wie hier aufklären, da dies die Ergebnisse beeinflussen könnte. Inzwischen haben Sie aber schon so viel über das Experimentieren gelernt, dass Sie als Experte dies ausblenden können. Um das Ganze nicht zu langweilig zu machen, wechseln wir wieder die Qualität. Es geht in diesem Versuch wieder um die Länge von Strichen. Lesen Sie also bitte jetzt gründlich die Instruktion!

10.2. Instruktion Versuch Strichtafel I und II

Bei diesem Versuch geht es darum, Striche unterschiedlicher Länge zu beurteilen. Auch hier werden auf jedem Bildschirm wieder die Pole gezeigt, also der kürzeste und der längste Strich. Dazwischen erscheint dann jeweils der Strich, den Sie beurteilen sollen. Benutzen Sie dazu eine Skala mit den drei Kategorien:

Kurz
Mittel
Lang

Wählen Sie die für diesen Strich zutreffende Kategorie aus. Es erscheint dann der nächste Strich. Urteilen Sie schnell und spontan und vergleichen Sie die Striche nicht miteinander!

Insgesamt werden Sie 60 Linien beurteilen, um wieder hinreichend viele Urteile zu haben. Bitte sorgen Sie jetzt wieder dafür, dass Sie ungestört arbeiten können und konzentrieren Sie sich auf diese Aufgabe!

Starten Sie danach das Programm Strichtafeln II. Es ist von der Aufgabe her völlig identisch. Allerdings wird dort eine andere Reizserie gegeben, wie Sie vermutlich schon vermutet haben. Lassen Sie sich durch dieses Wissen nicht beeinflussen und beurteilen Sie auch hier die Striche spontan und schnell.

Und hier geht es zum Versuch:

- <https://cgi-psych.rz.tu-bs.de>

10.3. Auswertung Strichtafeln I+II

Das Programm zeigt Ihnen pro Durchgang an, wie häufig Sie jede der drei Kategorien gewählt haben. Wir hatten im ersten Versuchsteil die 60 Striche ungefähr gleichabständig über den ausgewählten Längenbereich verteilt. Damit wäre in diesem Versuchsteil auch eine Gleichverteilung der Urteile zu erwarten. Da 60 Striche zu bewerten waren, wären bei einer Gleichverteilung $60/3 = 20$ Urteile in jeder Kategorie zu erwarten. Waren bei Ihnen die Urteile gleich häufig?

In einer Gruppe von 20 Studenten fanden sich die folgenden Häufigkeiten:

Kurz: 18

Mittel: 19

Lang: 23

Dies entspricht nicht völlig einer Gleichverteilung, ist ihr aber schon sehr nahe. Damit wurde in diesem ersten Versuchsteil gezeigt, dass auch bei der Beurteilung von Strichlängen eine annähernde Gleichverteilung zu finden ist. Ihnen ist vielleicht auch aufgefallen, dass hier ein Effekt der Pole nicht zu finden ist. Das ist auch durch unsere Kategorien zu erklären: „Kurz“ ist eben nicht der Pol, sondern ein ganzer Bereich von Strichlängen, ebenso wie „Lang“.

Im zweiten Versuchsteil waren kurze Striche häufiger vorhanden als längere Striche. Wenn man als Urteiler bestrebt ist, die verschiedenen Antwortkategorien gleich häufig zu benutzen, sollten sich die Ergebnisse des zweiten Versuchs nicht von dem ersten Versuch unterscheiden. Welche Häufigkeiten haben Sie in den drei Kategorien? Sind sie vergleichbar wie beim ersten Versuch?

Erstellen Sie dazu aus Ihren Daten eine Häufigkeitsverteilung, wie sie in Abbildung 41 dargestellt ist. Sie finden dazu in Excel eine Vorlage, bei der in der ersten Spalte die drei Kategorien stehen und in den folgenden beiden Spalten die Häufigkeiten der Urteile einzutragen sind. Die Abbildung zeigt sehr deutlich, dass sich die Verteilungen in den beiden Versuchsteilen unterscheiden. Wenn viele kurze Striche vorgegeben werden, findet man auch viele „kurz“-Urteile. Das ist an sich wenig verblüffend, aber durch diesen Befund können wir den Einwand ausräumen, dass der Effekt der Gleichverteilung eigentlich durch die Vorgabe der Kategorien entsteht. Das Ergebnis spricht vielmehr dafür, dass unsere Überlegungen zur Struktur von Bezugssystemen zutreffen: Der Bereich zwischen den Polen wird in gleichmäßig große Bereiche eingeteilt. Diese mentalen Kategorien werden jeweils durch einen Prototyp repräsentiert. Gibt man dann Reize, die gleichmäßig über dieses Kontinuum verteilt sind, so führen diese gleich großen Kategorien im Erleben dazu, dass auch die verwendeten Kategorien gleichhäufig genutzt werden.

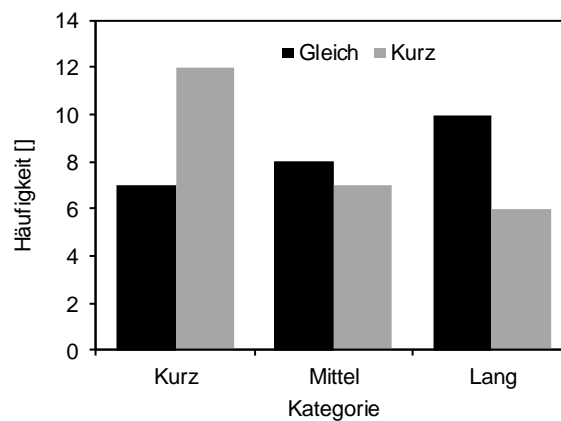


Abbildung 41: Häufigkeit der Kategoriennutzung in den beiden Versuchsteilung mit Gleichverteilung der Reize („gleich“) und bei häufigen kurzen Linien („kurz“). Dargestellt sind die Ergebnisse von 20 Probanden.

Wir können also festhalten:

- Wenn sich Bezugssysteme ausdifferenzieren, entstehen gleichabständige Kategorien, die über den Bereich möglicher Ausprägungen verteilt sind.
- Die Differenzierung heißt dann, dass die einzelnen Kategorien immer kleiner werden, so dass das Urteil immer genauer wird.

Dies hatten wir oben bereits in Abbildung 35 und Abbildung 36 so dargestellt, aber dort noch nicht beschrieben. Im ersten Schritt der Bildung eines Bezugssystems (Abbildung 35) gibt es nur die beiden Pole, also zwei sehr breite Kategorien. Dann entsteht die Mitte, wodurch das Kontinuum jetzt in drei Bereiche geteilt wird (Abbildung 36), die dann natürlich deutlich schmaler sind.

Mit diesem Versuch wollen wir unsere Studien zu Bezugssystemen abschließen. Das heißt natürlich nicht, dass damit schon alle Fragen zu Bezugssystemen geklärt sind. In den vorigen Kapiteln hatten wir jeweils im Text einige Fragen genannt, denen man noch nachgehen müsste. Aber das war auch nicht das Ziel des Buchs, dieses Thema vollständig zu behandeln. Vielmehr sollten Sie etwas darüber lernen, wie man als Psychologe methodisch vorgeht, um das Erleben zu untersuchen, also Dinge indirekt messbar zu machen, die nicht direkt gemessen werden können. Dass Sie dabei auch noch einiges über grundlegende psychische Prozesse bei der Wahrnehmung, über das Gedächtnis und über die Beurteilung von Objekten lernen, ist ein schöner Nebeneffekt. Im abschließenden Kapitel werden wir versuchen, die wichtigsten Dinge, die Sie lernen sollten, zusammenzufassen.

11. Fazit

Sie sind jetzt am Ende dieses Buches angekommen. Leider wartet kein weiterer Versuch auf Sie. Aber Sie müssen nicht traurig sein: Im Laufe des weiteren Studiums ergeben sich sicherlich noch einige Gelegenheiten, um eigene Studien durchzuführen.

Sie konnten im Rahmen dieser Übung eine Selbsterfahrung machen: Sie haben am Beispiel von Experimenten, die Sie im Selbstversuch durchgeführt haben, methodisch gelernt, dass Experimente **der** Weg der Psychologie sind, um Wissen zu erarbeiten und zu prüfen. Man stellt Theorien und Hypothesen auf und überlegt dann, wie man diese prüfen kann. Dazu überlegt man sich, was unter bestimmten Umständen passieren sollte, also wenn man z.B. sehr viele kurze Striche vorgibt. Sie werden im Rahmen der weiteren Veranstaltungen zur Methodenlehre noch weiter lernen, wie man gute Experimente plant und auswertet. Wir hoffen aber, dass Sie durch die verschiedenen Experimente im Selbstversuch dadurch schon sehr gut darauf vorbereitet sind.

Inhaltlich haben Sie etwas über ganz grundlegende psychische Prozesse, die Wahrnehmung, Beurteilung und Speicherung von Informationen, gelernt. Wie die Physik hat auch die Psychologie eine Metrik. Wir können als Menschen die Ausprägungen von bestimmten Eigenschaften erleben und in ein Bezugssystem einordnen. Dieses Bezugssystem, das unbewusst erlernt und verändert wird, ermöglicht uns absolute, zuverlässige Urteile. Ja vielmehr: indem wir ein Bezugssystem aufbauen, werden Objekte für uns erst zu Gegenständen mit Eigenschaften, die bei jedem konkreten Objekt in bestimmten Ausprägungen vorliegen. Unser ganzes Verständnis der Welt beginnt also mit dem Bezugssystem. Unser Langzeitwissen und unser Objektwissen ist im Gedächtnis als Menge von Bezugssystemen gespeichert.

Wir haben auch etwas über die Struktur von Bezugssystemen gelernt: Ganz wichtig sind die Pole, weil Sie den Bereich des Erlebens für bestimmte Objekte eingrenzen. Die Mitte des Bezugssystems (Adaptationsniveau) ist eine zusätzliche Beschreibung der Lage der Empfindungen. Zwischen den Polen besteht das Bezugssystem aus einer begrenzten Zahl von gleich großen Erlebensbereichen, die mental vermutlich in Form von Prototypen repräsentiert sind. Auf diese Weise können wir sehr differenziert und schnell neue Objekte in unserem Bezugssystem abbilden und so beurteilen.

Bezugssysteme entstehen durch Differenzierung. Mit zunehmender Erfahrung können wir Qualitäten immer feiner wahrnehmen. Mit zunehmender Erfahrung nimmt also die Anzahl dieser Kategorien zu, sodass Experten zuverlässiger und feiner urteilen können als Novizen. Wenn man so will, sind wir also ständig dabei, unsere Wahrnehmung zu erweitern – und das ganz ohne Drogen! Doch noch einmal kurz zurück zur Psychologie als Wissenschaft. In der Physik ist die Messung die fundamentale Methode, um die Gegenstände der Physik zu beschreiben. Ganz ähnlich ist es in der Psychologie, nur ist hier Messen etwas Anderes:

- Messen (in der Psychologie) = Einordnen in ein Bezugssystem

Die Psychologie will das Erleben und Verhalten messen. Das Verhalten kann man physikalisch messen – jemand reagiert in 250 ms, läuft den Marathon in 3:46 h usw. Um das Erleben zu messen, müssen wir auf das Bezugssystem zugreifen. Sie haben verschiedene Methoden kennengelernt, mit denen dies gelingen kann. In der Physik beschreiben wir den Gegenstand physikalisch. In der Psychologie geht es uns darum, wie Menschen die Welt (und sich selbst und andere Menschen) erleben. Wir hatten gelernt, dass sich ein solches Erleben als eine bestimmte (psychologische) Art des Messens beschreiben lässt. Erleben besteht darin, dass wir die erlebten Ausprägungen bestimmter Qualitäten in unser Bezugssystem einordnen und auf diese Weise dann die Quantität erleben. Dabei geschieht der erste Schritt der Einordnung unbewusst und

automatisch, so dass wir den Eindruck haben, unmittelbar zu erleben, wie eine bestimmte Eigenschaft ausgeprägt ist (absolutes Urteil).

Was nutzt uns dieses Wissen als Psychologen? Das Anliegen des Psychologen ist es doch, das Erleben von Menschen zu verstehen. Sie wissen jetzt, wie dies prinzipiell, von der Methode her, funktioniert. Damit sind aber natürlich Ihre konkreten Fragen bei einem konkreten Menschen oder einer konkreten Gruppe von Menschen noch nicht beantwortet. In diesem Sinne haben Sie vor allem etwas Methodik gelernt. Denken Sie an die Hinweise zurück, wie man gute Fragebögen erstellt! Da Sie jetzt wissen, wie unterschiedliches Erleben entsteht, sollten es für Sie sehr viel leichter sein, dies auch für Ihre eigenen psychologischen Fragestellungen zu nutzen. In diesem Sinne hoffen wir, dass wir Sie ein wenig für die Forschung interessieren und Ihr Verständnis für Methodik verbessern konnten.