

Abb. 4.9. Beispiel für die Rekonstruktion einer I-Skala (► Text)

Coombs-Skala

Dieser von Coombs (1948, 1950, 1952, 1953, 1964) entwickelte Skalentyp stellt die Untersuchungsteilnehmer vor die Aufgabe, eine Reihe von Items (z. B. Behauptungen), die unterschiedliche Ausprägungen des untersuchten Merkmals repräsentieren, nach Maßgabe ihres Zutreffens in eine Rangreihe zu bringen. Die individuelle Rangreihe ist nach diesem Ansatz von der Merkmalsausprägung der untersuchten Person bestimmt.

Nehmen wir an, man wolle das Stimulationsbedürfnis eines Untersuchungsteilnehmers *i* ermitteln. Dieser wird gebeten, die folgenden Items (nach Zuckerman et al., 1964) in eine Rangreihe zu bringen (die Beschränkung auf 4 Items dient nur der Vereinfachung der Demonstration):

- A. Ich gehe gern im Wald spazieren.
- B. Ich mag gemütliche Fahrten ins Blaue.
- C. Gelegentlich tue ich Dinge, die ein bißchen gefährlich sind.
- D. Ich würde gerne einmal selbst an einem Autorennen teilnehmen.

Die Rangreihe dieses Untersuchungsteilnehmers sei A, B, C, D, d. h., der Untersuchungsteilnehmer zieht offenbar beruhigende Tätigkeiten vor. Eine solche individuelle Rangreihe bezeichnet man als eine **I-Skala** (»Individual Scale«).

In Abb. 4.9 wird der (mögliche) »Stimulationsgehalt« der 4 Items verdeutlicht (genauer hierzu ► unten). Die Tatsache, dass Person *i* Item A auf Rangplatz 1 setzt, lässt darauf schließen, dass dessen Stimulationsgehalt dem Stimulationsbedürfnis der Person *i* am besten entspricht. Item A repräsentiert im Vergleich zu allen übrigen Items den »Idealpunkt« der Person *i*. Dementsprechend liegt die Position der Person *i* auf der Stimulationskala in der »Nähe« von A (Abb. 4.9). Eine solche Skala, die sowohl Items als auch Personen abbildet, bezeichnet man hier als **J-Skala** (»Joint Scale«).

Natürlich hätte man für Person *i* auch eine andere Position wählen können (z. B. links von A). Es muss

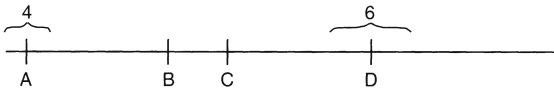
jedoch gewährleistet sein, dass die Distanz von *i* zu A kleiner ist als die von *i* zu B, denn sonst hätte die Person nach dieser Theorie B auf Rangplatz 1 und A auf Rangplatz 2 setzen müssen. Für einen anderen Untersuchungsteilnehmer, dessen Merkmalsausprägung wir mit 18 annehmen wollen, müsste die Rangreihe (I-Skala) lauten: C, B, D, A.

Für die Konstruktion einer J-Skala (und damit für die Ermittlung der Merkmalsausprägungen der untersuchten Personen und Items) aufgrund empirisch ermittelter I-Skalen hat nun der folgende Gedankengang zentrale Bedeutung: Ist das untersuchte Merkmal tatsächlich eindimensional und haben die Untersuchungsteilnehmer fehlerfreie (transitive, ► S. 160) Rangreihen abgegeben, existieren nur zwei Rangreihen, die zueinander spiegelbildlich sind. Diese Rangreihen stammen von Personen mit extremen Merkmalsausprägungen, die entweder in der Nähe (oder links) von A bzw. in der Nähe (bzw. rechts) von D liegen. Ihre Rangreihen müssten A, B, C, D bzw. – spiegelbildlich hierzu – D, C, B, A lauten. Eine dieser Rangreihen entspricht direkt der Rangfolge der Items auf dem Merkmalskontinuum. Alle übrigen Personenpositionen führen zu Rangreihen, für die es empirisch keine spiegelbildlichen Rangreihen geben darf, es sei denn, das Merkmal ist mehrdimensional oder die Rangreihen sind fehlerhaft. Für die Skalenkonstruktion ist es deshalb erforderlich, zwei zueinander spiegelbildliche Rangreihen zu finden.

Konstruktionsregeln. Wie die Skalenkonstruktion im einzelnen vor sich geht, sei im Folgenden an einem kleinen Beispiel demonstriert. Sieben Untersuchungsteilnehmer erhalten die Aufgabe, die oben genannten vier Behauptungen nach Maßgabe ihres Zutreffens in eine Rangreihe zu bringen. Sie nennen die folgenden Rangreihen (die Konstruktion der Skala folgt den Ausführungen von der Vens, 1980, S. 59 ff.):

1. C D B A
2. B C A D
3. C B D A
4. A B C D
5. C B A D
6. D C B A
7. B A C D

4.3 · Testen



■ **Abb. 4.10.** Vorläufige Positionen der Items A, B, C und D sowie der Personen 4 und 6

Unter den 7 Rangreihen befinden sich 2, die zueinander spiegelbildlich sind, und zwar die Rangreihen 4 (A, B, C, D) und 6 (D, C, B, A). Rangreihe 4 wird willkürlich als Rangfolge der 4 Behauptungen festgesetzt. (Rangreihe 6 würde zu einer J-Skala führen, die zu der hier zu entwickelnden J-Skala spiegelbildlich wäre.) Person 4 liegt offensichtlich in der Nähe von A und Person 6 in der Nähe von D (■ Abb. 4.10).

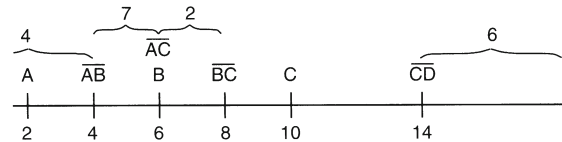
Die Abstände zwischen den Items sind hier zunächst beliebig; sie müssen lediglich den beiden spiegelbildlichen Rangfolgen genügen. Der Ausdruck »in der Nähe von« lässt sich nun insoweit präzisieren, als Person 4 auf jeden Fall näher an A als an B und Person 6 näher an D als an C liegen muß. Wählen wir als Skalenpunkte für A und B willkürlich die Werte 2 und 6, muss Person 4, die ja A vor B gesetzt hat, links vom Mittelwert \overline{AB} , also links von 4 liegen. Person 6 muss demzufolge rechts vom Mittelwert der Skalenwerte für C und D liegen. Dieser soll mit $\overline{CD} = 14$ angenommen werden. Es resultieren damit die folgenden Positionseinschränkungen für die Personen 4 und 6 (■ Abb. 4.11):

C und D sind vorläufig noch nicht bestimmt. C muss jedoch rechts von B liegen und links von \overline{CD} .

Als nächstes betrachten wir Personen, die Item B auf den ersten Rangplatz gesetzt haben. Es sind dies die Personen 2 und 7. Sie befinden sich offensichtlich in der Nähe von B oder genauer rechts von \overline{AB} und links von \overline{BC} . Der Mittelwert C wurde bereits auf 4 festgelegt. Der Mittelwert \overline{BC} muss links von \overline{CD} und natürlich rechts von B liegen. Diese Bedingungen sind erfüllt, wenn wir für \overline{BC} den Wert 8 annehmen. Da B bereits auf 6 festgelegt wurde, muss damit C den Wert 10 erhalten (■ Abb. 4.12).



■ **Abb. 4.11.** Positionen der Mittelpunkte \overline{AB} und \overline{CD}



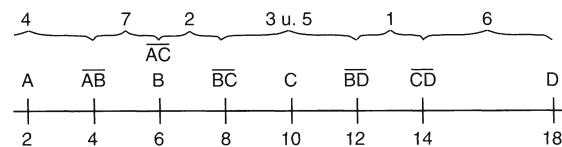
■ **Abb. 4.12.** Positionen der Mittelpunkte \overline{AB} , \overline{AC} , \overline{BC} und \overline{CD}

Der Unterschied zwischen den Personen 2 und 7 besteht in der Vergabe der Rangplätze 2 und 3 (Person 2: BCAD und Person 7: BACD). Person 2 liegt also näher bei C und Person 7 näher bei A oder: Person 2 befindet sich rechts von \overline{AC} und Person 7 links von \overline{AC} . Da A und C bereits festliegen (A=2, C=10), liegt auch \overline{AC} fest ($\overline{AC}=6$) (vgl. Abb. 4.12).

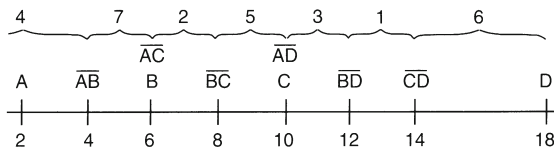
Die verbleibenden 3 Personen haben Item C auf Rangplatz 1 gesetzt, d. h., sie befinden sich in der Nähe von C bzw. rechts von \overline{BC} und links von \overline{CD} . Person 1 mit der Rangreihe CDBA setzt D vor B, d. h., sie liegt zusätzlich rechts von \overline{BD} . Für die Ermittlung von \overline{BD} gehen wir folgendermaßen vor: Wenn C den Wert 10 erhalten hat, und \overline{CD} auf 14 festgesetzt wurde, muss D den Wert 18 erhalten. Damit ergibt sich der Mittelwert \overline{BD} zu $(6+18):2=12$. Person 1 liegt zwischen den Werten 12(\overline{BD}) und 14(\overline{CD}) (■ Abb. 4.13).

Noch ungeklärt sind die genauen Positionen der Personen 3 (CBDA) und 5 (CBAD), die sich nur in den Rangplätzen 3 und 4 unterscheiden. Offensichtlich liegt Person 3 rechts von \overline{AD} (D wird A vorgezogen) und Person 5 links von \overline{AD} (A wird D vorgezogen). Für \overline{AD} ergibt sich mit $(2+18):2=10$ ein Wert, der – wie gefordert – zwischen \overline{BC} und \overline{CD} liegt. ■ Abb. 4.14 zeigt die Positionen aller Personen.

Aus den individuellen Rangreihen, die die Personen für die Items erstellen, lässt sich mit der hier beschriebenen Technik die Rangreihe der untersuchten Personen bezüglich des untersuchten Merkmals ableiten. Die Technik heißt nach Coombs **Unfolding Technique** (Entfaltungstechnik): »Faltet« man die J-Skala in einem Personenpunkt, geraten die links und rechts von diesem



■ **Abb. 4.13.** Positionen der Mittelpunkte \overline{AB} , \overline{AC} , \overline{BC} , \overline{CD} und \overline{BD}



■ **Abb. 4.14.** Positionen der Mittelpunkte \overline{AB} , \overline{AC} , \overline{BC} , \overline{CD} , \overline{BD} und \overline{AD}

Personenpunkt befindlichen Items auf eine Skalenseite. Ihre Rangfolge entspricht dann der Präferenzordnung der jeweiligen Person bzw. ihrer I-Skala. Coombs veranschaulicht das Unfolding anhand einer Schnur, auf der sich Knoten befinden, die die Positionen der Items und der Personen markieren. Wenn man nun die Schnur an einem Personenknoten ergreift und die Schnurende frei herunterhängen lässt, bildet die Abfolge der Itemknoten die Rangfolge der Items bzw. die I-Skala der betroffenen Person. Die J-Skala entsteht damit rückläufig durch Entfalten der einzelnen I-Skalen.

Die Konstruktion einer Coombs-Skala nach dem hier beschriebenen Verfahren ist bei größeren Item- und Personenzahlen sehr aufwendig. Man verwendet dann besser eine schematisierte Routine (die »Gleiche-Delta-Methode«), die z. B. bei van der Ven (1980, S. 66 ff.) ausführlich beschrieben wird.

Auf der Grundidee des »Unfolding« basierende Rechenprogramme sind im SAS-Programmpaket, im MDS-Programmpaket sowie in den Programmsystemen XGvis/XGobi (Swayne et al., 1998; ► Anhang D) enthalten.

Skaleneigenschaften. Hinsichtlich ihrer metrischen Eigenschaften entspricht die Coombs-Skala keiner der in ► Abschn. 2.3.6 behandelten Skalenarten. Die Positionen der Personenpunkte sind zwar nicht eindeutig festgelegt, können aber auch nicht beliebig variieren. Wir wissen nur, dass sich beispielsweise Person 7 zwischen den Mittelpunkten \overline{AB} und \overline{AC} befinden muss (■ Abb. 4.14). Eine präzisere Bestimmung der Personenpunkte ist nicht möglich. Derartige Bereiche, in denen die Personenpunkte frei variieren können, nennt Coombs »isotone Regionen«.

Für reine Ordinalskalen sind beliebige **monotone Transformationen** zulässig, also Transformationen, die die Rangordnung der untersuchten Objekte erhalten. Bei Coombs-Skalen ist hingegen darauf zu achten, dass

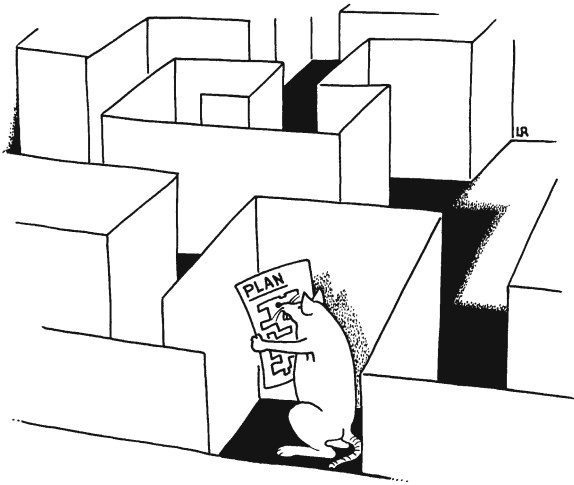
durch Transformationen die Rangfolge der Abstände zwischen Personen und Items bestehen bleibt (hypermonotone Transformation). Dieser Skalentyp, der bezüglich seiner Skalenqualität zwischen einer Ordinal- und einer Intervallskala anzusiedeln wäre, wird als **geordnete metrische Skala** bezeichnet.

Modellprüfung. Leider muss man in der Praxis häufig damit rechnen, dass nicht alle individuellen Rangreihen (I-Skalen) modellkonform und dass damit nicht alle untersuchten Personen skalierbar sind. Wie man sich leicht anhand ■ Abb. 4.14 überzeugen kann, wäre beispielsweise eine individuelle Rangreihe ADBC mit der gefundenen J-Skala nicht vereinbar. Insgesamt sind von $n!$ möglichen Rangreihen nur $0,5 \cdot n \cdot (n-1) + 1$ Rangreihen modellkonform. Für die vier im Beispiel verwendeten Items gibt es $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$ mögliche, aber nur $0,5 \cdot 4 \cdot 3 + 1 = 7$ zulässige Rangfolgen. Ist die Anzahl nicht zulässiger Rangreihen so groß, dass die Coombs-Skala praktisch unbrauchbar wird, können als Skalierungsalternativen eine von Bechtel (1968) entwickelte probabilistische Variante des »Unfolding« oder eine mehrdimensionale Unfoldingtechnik (Bennet & Hays, 1960; Hays & Bennet, 1961) verwendet werden.

Für die Entwicklung einer »klassischen« eindimensionalen Coombs-Skala ist es von Vorteil, wenn die »wahre« Rangordnung der verwendeten Items – eventuell aufgrund von Vorversuchen – bekannt ist oder doch zumindest theoretisch begründet werden kann. Ferner sollten die zu untersuchenden Personen möglichst das gesamte Merkmalspektrum repräsentieren, und das Merkmal selbst sollte aus der Sicht aller Personen eindimensional sein (vgl. hierzu auch Sixtl, 1967, S. 391 ff.).

Letztlich jedoch – und hierin ähnelt die Coombs-Skala der Guttman- oder auch der Rasch-Skala – führt die Entwicklung einer modellkonformen J-Skala nicht nur zu den Merkmalsausprägungen der untersuchten Personen, sondern auch zu einer Skalierung aller Items, d. h., auch die Coombs-Skala integriert den »Subject-centered«- und den »Stimulus-centered-Approach« (► S. 154).

Weitere Informationen zum »Unfolding« findet man z. B. bei Borg und Groenen (2005, Teil III), Borg und Staufenbiel (1993), Carroll (1983) sowie bei Rohwer und Pötter (2002, Kap. 15.3).



Testverfälschung: Wenn Probanden die Aufgaben nicht selber lösen. (Zeichnung: R. Löffler, Dinkelsbühl)

4.3.7 Testverfälschung

Testergebnisse, die nicht nur für wissenschaftliche Zwecke benötigt werden, haben für die getesteten Personen häufig lebenswichtige Konsequenzen. Sie entscheiden darüber, ob eine Abiturientin das Fach ihrer Wahl studieren darf, ob ein Arbeitnehmer den von ihm gewünschten Arbeitsplatz erhält, ob ein Schüler in eine Sonderschule eingeschult wird etc. Es ist deshalb keineswegs verwunderlich, wenn getestete Personen sich darum bemühen, ihre Testergebnisse in einer für sie möglichst günstigen Weise zu »korrigieren«. Negativ bewertete Aspekte ihrer Persönlichkeit werden verborgen und positiv angesehene überbetont oder erfunden (soziale Erwünschtheit), indem gezielt versucht wird, hohe Testwerte (Simulation) oder niedrige Punktzahlen (Dissimulation) zu erreichen. Die getesteten Personen können sich absichtlich verstellen und die Fragen aus der Perspektive einer von ihnen eingenommenen, fiktiven Rolle beantworten. In Leistungstests wird Wissen z. B. durch Raten (► S. 216 f.) simuliert; Dissimulation erreicht man durch »Dummstellen«.

Neben solchen absichtlichen Verfälschungen bzw. Verstellungen (»**Faking**«) können Test- und Fragebogenergebnisse auch von den Testpersonen unbemerkt und unkontrolliert verzerrt werden, weil besondere kognitive Effekte aus den Bereichen Gedächtnis, Konzentration,

Informationsverarbeitung, Selbstbeobachtung, Selbstdarstellung etc. auf die Testbeantwortung einwirken, sodass unaufmerksame, irrtümliche oder »zufällige« Ergebnisse resultieren. Auch die auf ► S. 183 ff. und ► S. 232 ff. dargestellten Urteilsfehler sowie Antworttendenzen (»**Response Sets**«), d. h., typische Reaktionen auf die Präsentation und Anordnung der Testaufgaben, sind unter Umständen gravierende Fehlerquellen. Zudem stellen Persönlichkeitstests oder Einstellungsfragebögen zuweilen Fragen, über die sich die untersuchten Personen bislang noch keine Gedanken gemacht haben und die deshalb mehr oder weniger beliebig beantwortet werden (vgl. hierzu auch Rorer, 1965).

Von Fehlern, Verzerrungen und Verfälschungen zu sprechen bedeutet, dass man implizit von der Existenz einer »wahren Merkmalsausprägung« bei der Testperson ausgeht, die sich möglichst unverfälscht im Testwert ausdrücken sollte und damit dem Testanwender hilft, sich hinsichtlich der interessierenden Merkmale ein genaues Bild von der getesteten Person zu verschaffen (zumindest ein genaueres Bild, als sich bei einer rein intuitiven Einschätzung ergeben würde). Der mit Testungen verbundene Aufwand ist stets an die Hoffnung geknüpft, sinnvolle, valide Informationen zu erhalten. Validitätseinbußen aufgrund von bewussten oder unwillkürlichen Antwortverzerrungen stellen den Wert einer Testung grundsätzlich in Frage; ihnen ist deswegen in der Methodenforschung viel Bedeutung beigemessen worden (z. B. Berg, 1967; Schwarz & Sudman, 1992).

Wie anfällig sind Tests für Verfälschungen? Es gibt praktisch keine Untersuchung, die nachweist, dass der jeweils geprüfte Test nicht verfälschbar wäre (vgl. hierzu eine Metaanalyse von Viswesvaran & Ones, 1999). Wenngleich noch nicht jeder Test auf seine Verfälschbarkeit hin untersucht wurde, muss man wohl davon ausgehen, dass die Verwertbarkeit von Testergebnissen generell von der Kooperationsbereitschaft der Testperson, der Zusammenstellung und Formulierung der Testitems sowie der Testsituation abhängt. Die meisten Untersuchungen zu dieser Thematik beschränken sich auf den Nachweis einer potenziellen Verfälschbarkeit von Testergebnissen. Wie stark welche Fehlerquellen in einer konkreten Untersuchung zu Buche schlagen, kann nicht allgemein vorausgesagt werden, sondern ist für jede einzelne Untersuchung genau abzuwägen und in

