

Die Frage war nun, wie man diese Parameter vermessen sollte. In seinem Test bezog sich Binet auf lebensnahe Aufgaben („Wie heißen die Monate? Wozu verwendet man einen Löffel?“) und staffelte die Fragen nach Schwierigkeit. Alle Fragen, die von 70 Prozent der Kinder eines Alters gelöst werden konnten, fasste er zu Reihen zusammen. Wenn ein Kind nun alle diese altersentsprechenden Fragen richtig beantworten konnte, waren Lebensalter und Intelligenzalter ident. Löste ein Kind auch Aufgaben der nächsthöheren Reihe, war das Intelligenzalter entsprechend höher.

Was in der Theorie einfach klang, war in der Praxis allerdings etwas komplexer: Idente Intelligenzalter-Unterschiede bezeichneten in unterschiedlichen Altersgruppen nicht idente Entwicklungsbeschleunigungen oder -verzögerungen, d.h. mit zunehmendem Alter

relativierte sich ein solcher Unterschied von beispielsweise zwei Jahren zunehmend. Überdies war sich Binet selbst nicht ganz sicher, was genau er da eigentlich vermaß, da die Aufgaben mithilfe sehr unterschiedlicher Fähigkeiten gelöst werden konnten.

Der Psychologe William Stern griff 1912 Binets Verfahren wieder auf und verbesserte es, indem er das Intelligenzalter am Lebensalter relativierte, der Psychologe Lewis M. Terman multiplizierte das Ergebnis mit 100. Stimmt Intelligenz- und Lebensalter überein, kam er mit dieser Methode auf 100, eine Zahl, die uns auch heute als Durchschnitts-IQ bekannt ist.

Die Ursprünge der Intelligenzmessung lagen also in der Untersuchung von Kindern und ließen sich nicht so ohne Weiteres auf Erwachsene übertragen. Selbst dem messwütigsten Untersucher war klar, dass die mit dem Alter fortschreitende Differenzierung

von Wissen und Fähigkeiten den Vergleich aller auf Basis von Wissensabfrage verunmöglichte, dass also ein Bauer über ganz andere Kenntnisse und Fertigkeiten verfügt als ein Atomphysiker und dass beide im jeweils anderen Beruf heillos überfordert wären, ohne deshalb als minder intelligent gelten zu müssen. 1930 wurde dieses Problem der Psychometrie von David Wechsler in einer heute noch gültigen Form gelöst. Wechsler, ein amerikanischer Psychologe rumänischer Abstammung (der eine psychoanalytische Ausbildung bei Anna Freud absolvierte), stellte bei der Untersuchung von Army-Rekruten fest, dass die gültige Intelligenz-Definition sehr zu hinterfragen war: Männer, die im Test erbärmlich schlechte Ergebnisse ablieferten, hatten keinerlei Probleme, ihr alltägliches Leben selbstbestimmt zu bewältigen. Wechsler kam daher zum Schluss, dass der Begriff

Intelligenz deutlich weiter gefasst werden musste und wohl auch nicht ausschließlich intellektuelle Anteile einer Person beinhalten sollte. Er definierte Intelligenz als „die zusammengesetzte oder globale Fähigkeit des Individuums, zweckvoll zu handeln, vernünftig zu denken und sich mit seiner Umgebung wirkungsvoll auseinanderzusetzen“.

Natürlich kann man auch diese aus vielen definitionsoffenen Begriffen zusammengesetzte Beschreibung hinterfragen: Zweckvoll für wen? Für den Betroffenen im Hier und Jetzt oder für eine größere Gruppe in näherer oder weiterer Zukunft? Und zu welchem Zweck? Was ist „vernünftig“? Was mir finanziell nützt? Was mir ohne finanziellen Vorteil Erfahrungen ermöglicht? Was mir finanzielle Nachteile bringt, aber mein Ansehen vermehrt? Und wie setze ich mich mit meiner Umgebung wirkungsvoll auseinander?

Fraglos kann man auch das Entleeren von Altöl in irgendeinen Abwasserkanal als wirkungsvoll ansehen, es kann für mich sogar zweckvoll sein, indem ich mich einer unbrauchbaren Substanz entledige, es kann mir vernünftig erscheinen, weil es mir weitere Wege oder gar Kosten spart, aber würden wir das als intelligente Handlung bezeichnen wollen?

Wechsler jedenfalls entwickelte einen Fragenkatalog, der unterschiedliche Fähigkeiten abprüfte und auch praktische Aufgaben umfasste, um das Übergewicht des verbalen Teils zu relativieren, und wandte ihn an einer großen Gruppe als Normierungsstichprobe an. Im Ergebnis zeigte sich die typische glockenartige Verteilung der Gauß'schen Verteilungskurve. Derjenige Wert, der am häufigsten gemessen wurde, wurde als „100“ definiert, alles darüber oder darunter war (bei einer Schwankungsbreite von +/- 15) also