

„Changes“ oder „Effort“? Eine Überprüfung von zwei Theorien zur Zeitschätzung*)

Joachim Funke

Psychologisches Institut der Universität Bonn

Theorien zur Zeitschätzung unterscheiden sich danach, welche Faktoren als einflußreich beim Zustandekommen von Zeittäuschungen erachtet werden. Kognitive Modelle nehmen etwa die Anzahl stattgefundenen Veränderungen („changes“) in einem bestimmten Zeitintervall oder das Ausmaß des im zu schätzenden Zeitintervall vorliegenden Verarbeitungsaufwands („processing effort“) als Determinanten an. In einem Entscheidungsexperiment sowie drei Replikationsstudien (jeweils $N = 1$) kann gezeigt werden, daß weder Veränderungen noch Aufwand die alleinigen Einflußgrößen darstellen. Im Unterschied zu anderen Arbeiten zeigt die vorliegende Untersuchung, wie möglicherweise ein Konfundierungsproblem gelöst werden kann, das bislang die unabhängige Variation der getesteten Determinanten der Zeitschätzung verhinderte. Zugleich weisen die Befunde auf große interindividuelle Unterschiede bei gleichzeitig hoher intraindividuellem Konsistenz hin. Dies belegt die Problematik der Aggregation von Daten über Individuen hinweg.

Subjektiv erlebte und objektiv verstrichene Zeit erweisen sich in vielen Alltagssituationen als diskrepant. Je nach Bedingung treten erhebliche „Täuschungen“ auf, die erklärungsbedürftig sind. Die vorliegende Arbeit untersucht zwei gängige Hypothesen, die das Zustandekommen von Zeitempfindungen und Zeittäuschungen beschreiben und erklären. Unter einer Zeittäuschung wird dabei die Abweichung der subjektiv empfundenen von der innerhalb eines begrenzten Intervalls tatsächlich vergangenen Zeit verstanden.

*) Den Studentinnen und Studenten Martin Fischer, Marlies Klein, Ralf Kretschmann und Claudia Main bin ich zu Dank verpflichtet. Im Rahmen eines experimentalpsychologischen Praktikums haben sie sich sowohl um die kritische Diskussion vorliegender Befunde als auch um die sorgfältige Durchführung der Studie verdient gemacht. Mein Kollege Edgar Erdfelder hat mir durch Anmerkungen und Kritik zu einer ersten Manuskriptfassung geholfen.

Problemstellung

Bevor auf die experimentelle Prüfung der Gültigkeit von Modellen der Zeitschätzung näher eingegangen wird, sollen zunächst einige allgemeinere Überlegungen zum Konzept der Zeitwahrnehmung und der Zeitempfindung vorangestellt sowie einige methodische Probleme bei empirischen Arbeiten in diesem Bereich erörtert werden.

Theorien und Befunde zur Zeitschätzung

Bisher erarbeitete theoretische Konzeptionen zur Zeitschätzung stammen aus der Physiologie bzw. der Kognitionspsychologie. Die physiologischen Modelle münden explizit in der Annahme, daß irgendwo im Organismus ein „Zeitorgan“ existiert, das quasi als innere Uhr fungiert (Aschoff, 1985). Allerdings ist es bisher noch nicht gelungen, ein derartiges Organ nachzuweisen, so daß diese Ansätze spekulativ bleiben (Borg & Galinat, 1985). Folglich wird in dieser Arbeit ausschließlich auf die kognitiv orientierten Modelle verwiesen. Diese untersuchen, inwiefern die erlebte Dauer eine Funktion der Reize ist, die während eines Zeitintervalls wahrgenommen, verarbeitet und gespeichert werden.

Der Unterscheidung von Fraisse (1984, S. 10) folgend wird in der vorliegenden Arbeit zugunsten der Zeitschätzung („estimation“) auf die Darstellung und Diskussion der reinen Zeitwahrnehmung („perception“) verzichtet, die sich auf Zeiträume innerhalb des „psychological present“, der wahrnehmbaren Gegenwart (James, 1890), d.h. etwa 2—3 Sekunden, beschränkt. Die hier untersuchten Intervalle sind größer als 5 Sekunden, so daß lediglich die Prozesse der „estimation“ relevant sind. Diese Schätzung wird nach Fraisse im Gegensatz zur „perception“ dadurch charakterisiert, daß zwei getrennte Ereignisse, von denen mindestens eines in der Vergangenheit liegt, miteinander verbunden und damit Gedächtnisprozesse in Anspruch genommen werden müssen (Fraisse, 1984, S. 9).

Vroon (1970, S. 115) läßt seinen Beitrag mit folgenden Worten beginnen: „The sense of time can not be studied without manipulating nontemporal variables.“ Dies bedeutet, daß Untersuchungen über Zeitwahrnehmung und Zeitschätzung im wesentlichen Untersuchungen zur Kontextabhängigkeit darstellen. Die Möglichkeiten zur Kontextvariation sind auch zugleich Möglichkeiten für die Ansatzpunkte verschiedener Theorien. Poynter und Homa (1983) folgen der von Block und Reed (1978) vorgeschlagenen Kategorisierung von Theorien zur Zeitschätzung, die auf der Art der Manipulationen am Inhalt eines zu schätzenden Zeitintervalls beruht. Unterschieden werden drei Klassen von Manipulationen:

(1) Manipulationen an der Anzahl intervenierender Ereignisse im zu schätzenden Zeitintervall

Eine darauf bezugnehmende Theorie ist z. B. das von Fraisse verfolgte Change-Modell (nachfolgend als „ChM“ abgekürzt). Bei ChM wird die empfundene Zeit abhängig gemacht von der Anzahl der wahrgenommenen Veränderungen („perceived changes“, vgl. Fraisse, 1984, S. 20). Auf eine einfache Formel gebracht kann man sagen: je mehr Veränderungen bei gleicher Gesamtzeit stattfinden, um so kürzer wird die subjektive Empfindung dieser Zeit. — Die Anzahl der Veränderungen kann z. B. durch die Anzahl von gestellten Aufgaben manipuliert werden.

(2) Manipulationen an Umfang oder Art der geforderten Informationsverarbeitung

Hierauf bezieht sich das etwa von Thomas und Weaver (1975) verfolgte „processing effort“-Modell („PEM“). Bei PEM ist der kognitive Aufwand, den jemand in einer gegebenen Zeit betreibt, für die Zeitempfindung ausschlaggebend. Nicht unterschlagen werden darf der Hinweis von Thomas und Weaver (1975, S. 363), wonach auch Speicher- und Abruf-Operationen die Zeitschätzung beeinflussen. Prinzipiell wird die Zeitschätzung als separate Aufgabe begriffen („temporal task“), für die je nach den sonstigen Anforderungen („non-temporal tasks“) wenig „Zeit“ bleibt. Auf eine einfache Formel gebracht kann man sagen: je höher die nichtzeitbezogenen Aufgabenanforderungen bei gleicher Gesamtzeit, um so kürzer die subjektive Zeitempfindung. Ergänzend ist zu bemerken, daß die als Stellvertreter für diesen Ansatz zitierten Autoren ihr Modell zunächst einmal für einen Zeitraum bis 100 msec formuliert haben. Eine Erweiterung des Gültigkeitsbereichs dieser Theorie auf einen darüber hinausreichenden Zeitraum wird von Fraisse (1984, S. 31) für äußerst sinnvoll erachtet und hier auch vollzogen. — Der „processing effort“ läßt sich durch unterschiedliche Schwierigkeiten der gestellten Aufgaben variieren.

(3) Manipulationen an der Komplexität des zu bearbeitenden Reizmaterials

Auf diese Größe bezieht sich z. B. das „storage size“-Modell („SSM“). Diese von Ornstein (1969) entwickelte und vertretene Theorie schreibt dem Gedächtnis eine zentrale Rolle zu, wobei als Variablen Anzahl und Komplexität der dargebotenen Reize definiert sind. Kritisch ist zu diesem Modell anzumerken, daß hinsichtlich der Operationalisierung der Variablen „Komplexität“ Schwierigkeiten auftreten, da diese sich entweder auf objektive Eigenschaften des Reizmaterials oder aber auf interne Gedächtnispro-

zesse des Probanden (Pb) beziehen kann. Eine empirische Überprüfung der Theorie ist auf eine Verknüpfung der „recall“- oder „recognition“-Leistung der Versuchsperson mit der Zeitschätzung angewiesen. „Mulligan and Schiffman (1979) and Ornstein (1969) demonstrated that providing a simplifying code for remembering (...) reduced estimates of the stimulus duration“ (Poynter & Homa, 1983, S. 549). Ebenfalls bestätigen die Befunde von Block (1974) den Verdacht einer experimentell nicht greifbaren unabhängigen Variablen: entweder als objektiv präsentierte oder als gespeicherte Zahl der Items. Weitere experimentelle Befunde, die mit SSM nicht in Einklang stehen, berichtet Block (1985).

Da sich also die individuelle Gedächtnisstrategie als bedeutsame intervenierende Variable nicht kontrollieren läßt und eine Reihe empirischer Befunde mit SSM nicht vereinbar sind, kann auf dessen Prüfung verzichtet werden. Wie lassen sich nun die beiden verbleibenden Modelle auf ihre Tauglichkeit hin untersuchen?

Konfundierungsproblematik und ein Lösungsvorschlag

Um ChM im Vergleich mit PEM experimentell prüfen zu können, muß man die nichtzeitbezogenen Aufgabenanforderungen unabhängig von der Anzahl der Veränderungen variieren. In bisherigen Studien waren die Variablen „changes“ und „processing effort“ konfundiert, d. h. nicht unabhängig voneinander variiert. Diese Konfundierung der unabhängigen Variablen macht die Aussagekraft von Experimenten jedoch angreifbar. Dies soll an einem Beispiel demonstriert werden.

In der Arbeit von Hicks, Miller und Kinsbourne (1976) wurde „processing effort“ realisiert als unterschiedliche Anforderung beim Sortieren von Spielkarten. Die Pbn mußten die verdeckt vorgelegten Spielkarten umdrehen und je nach Versuchsbedingung unterschiedlich „behandeln“: (1) unter einer Kontrollbedingung waren die Karten lediglich auf einen Stapel zu sortieren (= 0-bit-Bedingung); (2) die Karten waren nach den Farben „rot“ und „schwarz“ auf zwei Stapel aufzuteilen (= 1-bit-Bedingung); (3) die Karten waren nach der „Spiel“-Farbe (Kreuz, Pik, Herz, Karo) auf vier Stapel aufzuteilen (= 2-bit-Bedingung). Obwohl es den Autoren damit erfolgreich gelungen ist, den Verarbeitungsaufwand zu operationalisieren, macht ein im Nebensatz mitgeteilter Befund die Ergebnisse im Lichte des Vergleichs von ChM und PEM zweifelhaft: so konnte zwar gezeigt werden, daß im prospektiven Paradigma ein negativ linearer Zusammenhang zwischen Informationsgehalt der Aufgabe und geschätzter Zeit bestand (wobei die 1-bit-Gruppe die tatsächliche Zeitdauer fast korrekt schätzte), jedoch wurden mit wachsender Belastung weniger Karten bearbeitet. Dies bedeutet: Aufwand und Change waren negativ korreliert.

Nach diesen Überlegungen ist die Absicht, Effort und Change in einem Versuchsplan unabhängig voneinander zu variieren, bisher an der Tatsache gescheitert, daß bei *konstant* gehaltener Zeit keiner der beiden Faktoren variiert werden kann, ohne daß nicht auch der jeweils andere Faktor gegenständig betroffen wäre. Will man etwa den insgesamt zu erbringenden Effort konstant halten, aber die Zahl der Changes erhöhen, führt dies bei konstant gehaltener Intervalldauer zwangsläufig zu einer relativen Verkürzung der einzelnen Changes und damit wiederum zu einer Veränderung des Efforts. Will man umgekehrt die Zahl der Changes konstant halten, den Effort aber bei gleichem Zeitintervall erhöhen, resultiert eine relative Verminderung des Efforts.

Ein Ausweg aus diesem Dilemma besteht meines Erachtens darin, auf das konstante Zeitintervall zu verzichten und dem Pb soviel Zeit zur Verfügung zu stellen, wie er für die jeweiligen Efforts bzw. Changes benötigt. Die theoretischen Aussagen von ChM bzw. PEM sind ja nicht an absolute Größen geknüpft, sondern machen Aussagen über *relative* Veränderungen von Täuschungsbeträgen. Bei diesem Vorgehen wird in Kauf genommen, daß die unter verschiedenen Versuchsbedingungen verstreichende objektive Zeitstrecke variieren kann.

Methodische Überlegungen

Zwei wichtige methodische Aspekte sind zu berücksichtigen: (1) die Art der Versuchssituation und (2) die Art des verwendeten Urteilsverfahrens.

Hinsichtlich der Versuchssituation ist zwischen der prospektiven und der retrospektiven Methode der Zeitschätzung zu unterscheiden. Im *prospektiven Paradigma* wird der Pb vor Versuchsbeginn darüber unterrichtet, daß er die Dauer von Zeitintervallen schätzen soll. Brown (1985) nimmt an, daß der Pb dadurch besonders motiviert sei, das Verstreichen der Zeit zu überwachen und brauchbaren Anhaltspunkten Aufmerksamkeit zu schenken, um eine möglichst präzise Zeitschätzung erbringen zu können.

Im Gegensatz dazu wird der Pb im *retrospektiven Paradigma* unvorbereitet zur Schätzung des gerade verstrichenen Zeitintervalls aufgefordert. In diesem Fall vermutet Brown (1985), daß der Pb die Informationen über die vergangene Zeitspanne eher zufällig und damit unzuverlässiger als im prospektiven Paradigma verarbeitet.

Die Experimente, die Brown (1985) in Anlehnung an das PEM nach Thomas und Weaver (1975) unter prospektiver und retrospektiver Bedingung durchgeführt hat, ergaben, daß die Urteile der Pbn im prospektiven Paradigma länger und korrekter ausfallen als die der Pbn im retrospektiven Paradigma. Fraise (1984) weist auf entsprechende Ergebnisse von Miller, Hicks und Willette (1978) hin.

Die grundlegende Ähnlichkeit prospektiver und retrospektiver Zeitschätzung zeigt sich jedoch bei steigender Aufgabenschwierigkeit. In beiden Paradigmen waren die Urteile kürzer und/oder ungenauer, wenn die gestellte Aufgabe schwieriger wurde: „As nontemporal processing demands increase, less capacity is available for processing temporal cues, and a consequent decrease or unreliability in duration judgments is observed.“ (Brown, 1985, S. 119). Der Effekt tritt im prospektiven Paradigma deutlicher hervor: unter dieser Bedingung sei der Pb motiviert, Zeithinweise aktiv zu suchen, während er sich im retrospektiven Paradigma auf die noch erinnerbaren Informationen verlassen muß. — In dem hier dargestellten Experiment wird ausschließlich mit der prospektiven Methode gearbeitet, da mit dem retrospektiven Paradigma von jedem Pb nur ein einziges Urteil erfaßt werden kann. Theorien mit Individuenparameter lassen sich so nicht testen.

Zur *Erfassung der wahrgenommenen Zeit* gibt es drei wesentliche Methoden: verbale Schätzung, Produktion und Reproduktion (vgl. Bindra & Waksberg, 1956; Hornstein & Rotter, 1969). Diese Urteilsverfahren unterscheiden sich bezüglich der Tätigkeiten von Versuchsleitern (Vl) und Pbn.

Bei der *verbalen Schätzung* begrenzt der Vl operativ ein Intervall und erfragt anschließend vom Pb dessen Dauer in Sekunden. Im Falle der *Produktionsmethode* gibt der Vl ein Intervall verbal vor und instruiert den Pbn, dieses operativ herzustellen. Wenn zunächst der Vl und daran orientiert der Pb operativ ein Intervall herstellen, liegt die sogenannte *Reproduktionsmethode* vor. Weitere Methoden lassen sich als Varianten dieser genannten Verfahren auffassen (Hornstein & Rotter, 1969, S. 561).

Die Besonderheit der in der vorliegenden Arbeit benutzten Variante der Reproduktionsmethode besteht darin, daß die Pbn sowohl das zu beurteilende Intervall als auch dessen Reproduktion selbst operativ herstellen. Fraise weist darauf hin, daß mit dieser Methode kurze Zeiträume bis zu einigen Minuten sehr gut exploriert werden können: „Die auf der Reproduktionsmethode basierenden Untersuchungen sind am zuverlässigsten...“ (Fraise, 1985, S. 215).

Hypothesen

Um angesichts der geschilderten konkurrierenden Theorien zur Zeitschätzung in dem Bereich größer als 10 Sekunden zu einem klaren Ergebnis zu kommen, werden zunächst die wissenschaftlichen Hypothesen (wH) aus der Sicht der beiden zu prüfenden Theorien, ChM und PEM, geschildert. Dabei wird deutlich, daß unter sonst gleichen experimentellen Bedingungen andere Prognosen resultieren. Zugleich erfolgt die Umsetzung der wissenschaftlichen in statistische Hypothesen (sH).

Bei unabhängiger Variation von „Change“ und „Effort“ in einem faktoriellen Versuchsplan sind verschiedene Möglichkeiten und damit verbunden unterschiedliche Prognosen denkbar (alle Aussagen beziehen sich auf eine prospektive Datenerhebungssituation):

wH (1):

Bei gleicher Anzahl von „Changes“, aber unterschiedlichem „Effort“ würde nach ChM keine Veränderung der Zeitschätzung erwartet, nach PEM aber mit wachsendem „Effort“ eine wachsende Unterschätzung eintreten.

wH (2):

Bei gleichem „Effort“, aber unterschiedlichen „Changes“ würde nach PEM nichts passieren, nach dem ChM jedoch mit zunehmenden „Changes“ eine wachsende Unterschätzung eintreten.

Die verschiedenen Effekte, die je nach Standpunkt zugelassen bzw. verboten werden, lassen sich als einfache Haupteffekte beschreiben, wobei die Mittelwerte benachbarter Faktorstufen jeweils einem zumindest monotonen Trend folgen sollten.

Für die wH (1) des PEM läßt sich somit folgende Konjunktion statistischer Hypothesen angeben, deren Zutreffen mit den theoretischen Erwartungen vereinbar wäre:

sH (1):

Kein Haupteffekt „Changes“, keine Interaktion beider Faktoren *und* positiver monotoner Trend zwischen den jeweils benachbarten Stufen des Faktors „Effort“.

Analog läßt sich die wH (2) des ChM in folgende Konjunktion statistischer Hypothesen umsetzen:

sH (2):

Kein Haupteffekt „Effort“, keine Interaktion beider Faktoren *und* positiver monotoner Trend zwischen den jeweils benachbarten Stufen des Faktors „Changes“.

Wie man unschwer erkennt, schließt jedes der beiden Modelle den jeweils anderen Haupteffekt aus und behauptet für sich einen positiven monotonen Trend. Beide Modelle verbieten das Auftreten einer Interaktion. Bei derartigen Hypothesenlage sollte eine klare Entscheidung zugunsten eines der beiden oder gegen beide Modelle möglich sein.

Methode

Versuchsplan

Als *unabhängige Variablen* sollen somit die zwei Faktoren „Effort“ und „Changes“ als „fixed factors“ mit je mehreren Stufen realisiert werden.

Analog zum Experiment von Hicks, Miller und Kinsbourne (1976) wird die Belastung über die Art der Aufgabe beim Sortieren von Spielkartenstapeln operationalisiert, während sich die Veränderungen aus der Anzahl der zu bearbeitenden Spielkarten ergeben. Um hier einen breiteren Bereich als üblich abzudecken, soll der Faktor „Changes“ mit den Stufen 8, 16, 32, 64 und 128 zu sortierende Karten hergestellt werden. Beim Faktor „Effort“ wird die Sortieraufgabe darin bestehen, die Karten in 1, 2, 4 oder 8 Stapel zu ordnen. Die Karten sind je nach Bedingung (1) nur in ein Fach zu stecken, (2) nach den Farben „rot“ oder „schwarz“, (3) nach den Spielfarben „Kreuz“, „Pik“, „Herz“, „Karo“, oder (4) nach dem Kartenwert „7“, „8“, „9“, „10“, „Bube“, „Dame“, „König“, „As“ zu sortieren. Dies entspricht einem zu verarbeitenden Informationsgehalt von 0, 1, 2 und 3 bit pro Item. Damit resultiert ein vollständig gekreuzter zweifaktorieller 4×5 -Versuchsplan mit 20 Zellen.

Als *abhängige Variable* dient das logarithmierte Verhältnis der subjektiv empfundenen zur tatsächlich vergangenen Zeit, also ein relatives Maß.¹⁾ Dies ist in bezug auf die vorgestellten theoretischen Formulierungen eine angemessenere Variable als die absolute Höhe der geschätzten Zeit. Ähnliche Vorgehensweisen sind in Untersuchungen zu geometrisch-optischen Täuschungen zu finden (vgl. Bredenkamp, 1984), in denen z. B. nicht die geschätzte Streckenlänge, sondern ihr Verhältnis zur objektiv gegebenen Länge interessiert. Eine Konsequenz derartiger AVn ist es, daß die aufgeklärte Varianz der Täuschungsbeträge wesentlich niedriger ausfällt als die der absoluten Urteile. Dies reflektiert aber lediglich den trivialen Sachverhalt, daß die geschätzte Dauer natürlich zum größten Teil von der tatsächlich verstrichenen Zeit abhängt.

Das *Urteilsverfahren*, nach dem ein Pb die beim Bearbeiten einer Aufgabe vergangene Zeit angeben soll, besteht zum einen in der Reproduktionsmethode, zum anderen in einer anschließenden verbalen Angabe der geschätzten Zeit in Sekunden. Somit resultieren zwei verschiedene AVn, die jeweils als Logarithmus des Täuschungsbetrages bei der Reproduktion (log TR) und der verbalen Schätzung (log TV) bestimmt werden.

Um die genannten experimentellen Bedingungen herstellen zu können, wird als *Versuchsmaterial* auf Skatkarten zurückgegriffen. Dort finden sich die Angaben, nach denen die oben beschriebene unterschiedlich differenzierte Sortierung erfolgen kann. Insgesamt zehn Sätze zu je 32 Karten eines bekannten Skatkarten-Produzenten fanden Verwendung.

1) In den genannten Theorien finden sich keine Hinweise darauf, welche Operationalisierung des Täuschungsausmaßes angemessen ist. Denkbar wäre etwa auch, Differenzen zwischen objektiv und subjektiv verstrichener Zeit zu bestimmen. Diese würden zu anderen Rangordnungen führen als die verwendeten relativen Täuschungsbeträge.

Hinsichtlich der *Anzahl der Beobachtungen* pro Zelle des Versuchsplans wird davon ausgegangen, daß es sich bei den postulierten Effekten um große Effekte im Sinne Cohen's (1977) handelt. Die wissenschaftlichen Hypothesen von ChM und PEM werden in statistische Hypothesen übersetzt, die als Konjunktion jeweils signifikanter Mittelwertsunterschiede benachbarter Faktorstufen bei gleichzeitiger Nicht-Signifikanz des jeweils anderen Faktors darstellbar sind. Am Beispiel des ChM bedeutet dies: seine Gültigkeit impliziert, daß die mittleren Täuschungsbeträge mit zunehmender Zahl von Veränderungen wachsen (geprüft via t-Tests zwischen jeweils benachbarten Mittelwerten), also ein positiver monotoner Trend des Change-Faktors vorliegt, und *gleichzeitig* kein bedeutsamer Effekt des anderen Faktors eintritt (geprüft via pauschalem F-Test). Da die beiden wissenschaftlichen Hypothesen jeweils einer Konjunktion statistischer Hypothesen entsprechen, die die Gültigkeit von einer H_0 und einer H_1 implizieren, muß eine entsprechende Korrektur der Irrtumswahrscheinlichkeit für den gesamten Test vorgenommen werden (vgl. hierzu genauer Westermann & Hager, 1986). Bei einem akzeptierten Alpha- wie Beta-Risiko von 0.10 sowie einer erwarteten Effektstärke von $f = 0.40$ für den jeweiligen F-Test und von $d = 0.80$ für die jeweiligen t-Tests resultiert ein notwendiges N zwischen 7 und 8 Beobachtungen pro Zelle. Die größere Zahl wurde ausgewählt, so daß für die 20 Versuchszellen jeweils 8 Messungen vorzunehmen waren, insgesamt also 160 Daten vorliegen sollten.

Die Daten wurden als *Meßwiederholungen* an einer Person erhoben. Der Grund für dieses Vorgehen ist darin zu sehen, daß die Theorien als Aussagen über Individuen zu verstehen sind und keinen Hinweis darauf geben, daß die Effekte nur auf Aggregat-Ebene in Erscheinung träten — im Gegenteil: vielfach werden Redensarten als unterstreichende Aussagen herangezogen (so z. B. bei Fraisse, 1984, S. 23: „A watched pot never boils“), die ja gerade die Allgemeingültigkeit hervorheben. Auf der anderen Seite ist bekannt, daß es eine Reihe von Individuen-Parametern gibt, die die Zeitschätzung beeinflussen (zur Übersicht siehe Fraisse, 1984, S. 25–26). Da die vorliegenden Theorien im Sinne universeller Hypothesen zu interpretieren sind, stellt die Einzelfallanalyse in diesem Fall „die strengstmögliche Prüfung“ (Leichsenring, 1987, S. 105) dar, sofern man durch versuchstechnische Vorkehrungen Störungen der internen Validität des Experiments durch Größen wie Training, Ermüdung, usw., ausschließen kann (vgl. die entsprechenden Warnungen bei Hager, 1987, S. 109). Der Vorteil dieses Verfahrens besteht darin, daß man die individuellen Parameter von der Form einer möglichen Gesetzmäßigkeit des Zusammenhangs zwischen der Zeitschätzung und den experimentellen Bedingungen trennen kann, wenn man dieses Experiment an mehreren Pbn repliziert. Zugleich erhöht dieses Vorgehen die Strenge der Prüfung.

Versuchsdurchführung

Die Datenerhebung fand im Frühjahr 1987 in einem 16 qm großen Kellerraum des Psychologischen Instituts statt, der durch eine Stellwand in einen Teil für den Vl und einen für den Pb unterteilt war. Im Raum befanden sich keine Uhren, Armbanduhren mußten abgelegt werden. Zwei weibliche und zwei männliche Vl betreuten in unterschiedlichen Paarungen jeweils einen Pbn, wobei der erste Vl für die Zeitmessung, der zweite Vl für die Vorbereitung der verschiedenen Bedingungen verantwortlich war. Der Pb las in der ersten von jeweils mehreren Sitzungen die Instruktion, in der er die Aufgabenstellung — Sortierung von Spielkarten nach unterschiedlichen Kriterien — und den Versuchsablauf beschrieben fand. Zugleich wurde er gebeten, während des Sortierens nicht zu zählen. Auch das Verfahren zur Reproduktion und zur Abgabe einer verbalen Schätzung wurden vorgestellt. Während der gesamten Versuchsdurchführung trug der Pb einen Kopfhörer als Schallschutz, während der Reproduktionsphase zusätzlich eine Schlafbrille als Sichtschutz; letztgenannte Maßnahme wurde von zwei Pbn durch Augenschließen ersetzt.

Vor dem Pb befand sich ein Kartenspender²⁾, der die zu sortierenden Skatkarten enthielt, ohne daß der Pb einen Anhaltspunkt für den noch verbleibenden Kartenumfang besaß. Die jeweils mit dem Rücken sichtbare Karte war dem Spender zu entnehmen und je nach Sortierbedingung zu bearbeiten. Hierzu befand sich ebenfalls vor dem Pb ein Sortierkasten (Frontfläche ca. 40×40 cm), auf dessen Vorderseite zwei mal vier untereinander angeordnete rechteckige Sortieröffnungen (ca. 11×3 cm) angeordnet waren; die Menge der bereits sortierten Karten war durch diese Anordnung nicht abzulesen. Je nach Verarbeitungsaufwand wurde eine Schablone vor diese Front gehängt, die nur eine Öffnung frei ließ (rechts unten) bzw. die unteren zwei, die unteren vier oder alle Öffnungen freigab. Zur Identifikation der Öffnungen waren je nach Bedingung entsprechende Symbole angebracht.

Jeder Durchgang wurde vom Pb durch Drücken einer Start-Taste eingeleitet, die eine hinter der Stellwand befindliche Stoppuhr startete. Der Pb warf dann die aus dem Kartenspender gezogene und verbal identifizierte Karte entsprechend dem Sortierkriterium in die zugehörige Öffnung. Er zog daraufhin die nächste Karte solange, bis eine „Stop“-Karte das Ende der Bearbeitung signalisierte. Zugleich betätigte er eine Stop-Taste. Den Beginn der daran anschließenden Reproduktion zeigte der Pb durch erneuten Druck der Start-Taste an; nochmalige Tastenbetätigung beendete die Reproduktion. Im Anschluß daran gab der Pb seine verbale Schätzung des verstrichenen Zeitintervalls an. Während der Reproduktion sollte der Pb nicht sprechen und sich möglichst wenig bewegen, um Verzerrungen durch externe Reize so gering wie möglich halten. Jeden Tastendruck gab der Pb mit den Worten „Start“ bzw. „Stop“ an.

Die Versuche wurden zwischen 8 und 21 Uhr durchgeführt, je nach Terminabsprache über einen Zeitraum von drei bis sieben Tagen. Die Dauer eines Durchgangs (20 Bedingungskombinationen) betrug zwischen 50 und 120 Minuten. Die Bedingungen folgten zufällig aufeinander. Sobald alle im Kartenspender befindlichen Karten sortiert waren (maximal 320 Karten), entstand eine Pause von 5 bis 20 Minuten, in der die Vl die nächsten Bedingungskombinationen vorbereiteten. Pro Sitzung sollte jeder Pb mindestens einen Durchgang bearbeiten, tatsächlich wurden zwischen 10 und 90 Sortierbedingungen hintereinander bearbeitet. Nach acht Durchgängen war das Experiment beendet.

2) Ich danke dem Landeskriminalamt Nordrhein-Westfalen in Düsseldorf für die vorübergehende Bereitstellung eines Kartenschlittens, wie er normalerweise in Spielsalons zum Einsatz kommt.

Probanden

Insgesamt vier Pbn nahmen an der Untersuchung teil. Bei den Pbn „Birg“ (22jährig, weiblich), „Mark“ (23jährig, männlich) und „Mich“ (28jährig, männlich) handelt es sich um Psychologiestudenten; diese meldeten sich freiwillig auf eine Nachfrage in einer Veranstaltung. Sie erhielten für ihre Mitarbeit einen im Rahmen des Grundstudiums notwendigen Nachweis über die Ableistung entsprechender Vpn-Stunden.

Beim vierten Pb, „Joac“ (34jährig, männlich), handelt es sich um den Autor dieses Berichts, dessen Daten mit entsprechender Vorsicht (Kenntnis der Hypothesen) zu interpretieren sind. Andererseits konnten hier wichtige Details der Versuchsdurchführung (kein Zählen während der Bearbeitung; konzentrierte Mitarbeit) sichergestellt werden.

Ergebnisse

Hypothesenbezogene Befunde

Zunächst interessieren die Resultate der hypothesenbezogenen Datenanalyse im Hinblick auf die Vereinbarkeit der beiden Modelle mit den Daten. Tabelle 1 zeigt die Ergebnisse der pauschalen F-Tests für Haupteffekte und Interaktion.

Das Ergebnismuster ist zwar nicht ganz eindeutig, aber bis auf zwei Fälle lassen sich die Daten schon nach den pauschalen F-Tests mit keinem der beiden Modelle vereinbaren. Es liegen zwar signifikante Effekte vor, diese treten jedoch hypothesenkonträr bei beiden Faktoren gleichzeitig auf.

Betrachtet man die beiden modellverträglichen Fälle — log TV bei Pb Mark und log TR bei Pb Mich — nun hinsichtlich des zweiten Teils der konjunktiven sH genauer, zeigt ein Blick auf die in Tabelle 2 mitgeteilten und in Abbildung 1 veranschaulichten Zellenmittelwerte bei Pb Mark den Verstoß gegen die postulierte Monotonie.

Einzig für Pb Mich ist das Vorzeichenmuster der t-Tests bei log TR hypothesenkonform, aber leider erreicht nur der Kontrast der Stufe 32 mit 64 die erforderliche Signifikanz. Würde man die statistische Hypothese derart liberalisieren, daß benachbarte Mittelwerte größer oder gleich sein dürfen, wären diese Daten damit verträglich.

Die Wirkung beider Faktoren geht im übrigen in die jeweils prognostizierte Richtung (vgl. Abbildung 1), wobei diese Effekte bei log TR deutlicher zutage treten als bei log TV. Mit wachsendem Effort verkürzt sich in fast allen Fällen die wahrgenommene Zeit, d. h. die Überschätzung nimmt ab (Pb Birg) bzw. die Unterschätzungen nehmen zu (Pbn Mark, Mich,

Tabelle 1

Ergebnisse der F-Tests auf Haupteffekte „Changes“ (CH) und „Effort“ (PE) sowie deren Interaktion getrennt nach zwei abhängigen Variablen und nach den vier Pbn. Die Spalten H(CHM) und H(PEM) geben an, ob die Daten mit einem der beiden Modelle verträglich sind.

AV*	F-Bruch des Effekts			mult. R ²	H(CHM)	H(PEM)
	CH	PE	CH*PE			
			Birg			
log(TR)	59.46*	6.49*	2.03*	0.67	nein	nein
log(TV)	77.57*	6.88*	0.68	0.70	nein	nein
			Mark			
log(TR)	6.94*	3.30*	0.60	0.24	nein	nein
log(TV)	4.86*	1.37	0.86	0.19	ja	nein
			Mich			
log(TR)	11.31*	1.94	1.04	0.31	ja	nein
log(TV)	1.64	1.17	0.63	0.11	nein	nein
			Joac			
log(TR)	31.62*	3.64*	1.19	0.52	nein	nein
log(TV)	0.38	0.36	1.07	0.10	nein	nein

*log TR = logarithmierter Täuschungsbetrag bei Reproduktion; log TV = logarithmierter Täuschungsbetrag bei Verbalschätzung.

*p ≤ 0.10 bei df_{CH} = 3, df_{PE} = 4, df_{CH*PE} = 12 und df_{Error} = 140.

Joac). Gleiches gilt für zunehmende Changes. — Tabelle 2 enthält im unteren Drittel auch die mittleren objektiven Sortierzeiten unter den verschiedenen Versuchsbedingungen. Daran ist — triviale — Tatsache abzulesen, daß die jeweiligen Treatments Effekte in die erwartete Richtung hatten: mit wachsender Belastung steigt die Sortierzeit ebenso wie mit wachsender Kartenzahl. Varianzanalysen mit der AV „log OBZ“ weisen für alle vier Pbn jeweils signifikante Haupteffekte bei gleichzeitig ausbleibendem Interaktionseffekt nach.

Um der Frage nachzugehen, inwiefern die Varianz der Täuschungsbeträge von der unter den verschiedenen Bedingungen variablen Zeitdauer abhängt, zeigt Tabelle 3 die Ergebnisse der F-Tests unter Einbezug der Kovariate „logarithmierte objektive Zeit“ (log OBZ).³⁾

3) Die Kovarianzanalyse folgt einem Vorschlag eines Reviewers. Es ist jedoch fraglich, inwiefern die für eine derartige Analyse u. a. notwendige Voraussetzung erfüllt ist, wonach die Kovariate der UV kausal vorgeordnet sein sollte (vgl. Bredenkamp, 1980, S. 79). Im vorliegenden Fall ist die Kovariate LOGOBZ den Treatment-Bedingungen kausal nachgeordnet.

Tabelle 2
Mittelwerte und Streuungen (M/SD) für die jeweiligen Faktorstufen von Effort (PE) und Changes (CH), getrennt nach den vier Pbn und für zwei verschiedene Täuschungsmaße sowie für die objektive Sortierzeit.

Zelle	Pb			
	Birg	Mark	Mich	Joac
	log TR (Reproduktion) ^a			
PE = 1	0.48/0.64	-0.43/0.26	-0.48/0.40	-0.40/0.34
PE = 2	0.27/0.51	-0.46/0.30	-0.53/0.41	-0.51/0.38
PE = 3	0.23/0.45	-0.50/0.27	-0.67/0.38	-0.56/0.35
PE = 4	0.16/0.59	-0.61/0.32	-0.58/0.42	-0.60/0.32
CH = 1	0.88/0.48	-0.29/0.29	-0.33/0.21	-0.39/0.23
CH = 2	0.52/0.33	-0.56/0.26	-0.41/0.33	-0.28/0.25
CH = 3	0.35/0.35	-0.46/0.30	-0.53/0.33	-0.37/0.21
CH = 4	0.02/0.38	-0.58/0.26	-0.71/0.41	-0.62/0.32
CH = 5	-0.34/0.31	-0.61/0.25	-0.85/0.48	-0.93/0.30
	log TV (verbale Schätzung) ^a			
PE = 1	-0.19/0.50	-0.77/0.28	-0.88/0.38	-0.73/0.26
PE = 2	-0.35/0.42	-0.85/0.31	-0.94/0.45	-0.75/0.24
PE = 3	-0.40/0.40	-0.82/0.29	-1.04/0.40	-0.76/0.33
PE = 4	-0.44/0.48	-0.91/0.38	-0.92/0.32	-0.80/0.40
CH = 1	0.18/0.28	-0.74/0.29	-0.85/0.28	-0.79/0.37
CH = 2	-0.12/0.29	-0.98/0.27	-0.88/0.34	-0.74/0.31
CH = 3	-0.34/0.29	-0.85/0.35	-0.92/0.39	-0.73/0.24
CH = 4	-0.56/0.34	-0.91/0.26	-1.05/0.44	-0.81/0.31
CH = 5	-0.88/0.17	-0.70/0.34	-1.03/0.45	-0.73/0.30
	OBZ (Sortierzeit) ^a			
PE = 1	74.0/73.0	63.7/56.6	71.6/60.6	66.8/59.3
PE = 2	79.0/71.7	78.5/68.1	83.7/73.8	76.7/68.4
PE = 3	90.1/75.9	87.9/78.3	93.3/80.2	86.9/73.0
PE = 4	98.9/89.3	95.9/83.2	100.1/88.0	92.1/81.5
CH = 1	15.1/ 2.4	16.1/ 2.7	16.6/ 2.6	15.0/ 2.3
CH = 2	29.3/ 5.5	29.5/ 6.7	31.9/ 4.2	27.7/ 5.1
CH = 3	55.5/10.7	52.2/10.0	58.7/ 9.2	52.3/ 7.5
CH = 4	105.2/26.5	100.9/19.7	112.6/24.3	103.9/13.2
CH = 5	222.2/33.1	208.7/36.2	216.0/47.6	204.7/32.0

^a log TR = logarithmierter Täuschungsbetrag bei Reproduktion; log TV = logarithmierter Täuschungsbetrag bei Verbalschätzung; OBZ = objektive Sortierzeit.

Vergleicht man diese Resultate mit den in Tabelle 1 berichteten Befunden, ist in drei von acht Fällen ein Ergebnismuster zu konstatieren, das nach dem Effort-Modell zu erwarten war. In keinem einzigen Fall kann ein mit dem Change-Modell verträgliches Ergebnis festgestellt werden. Bis auf

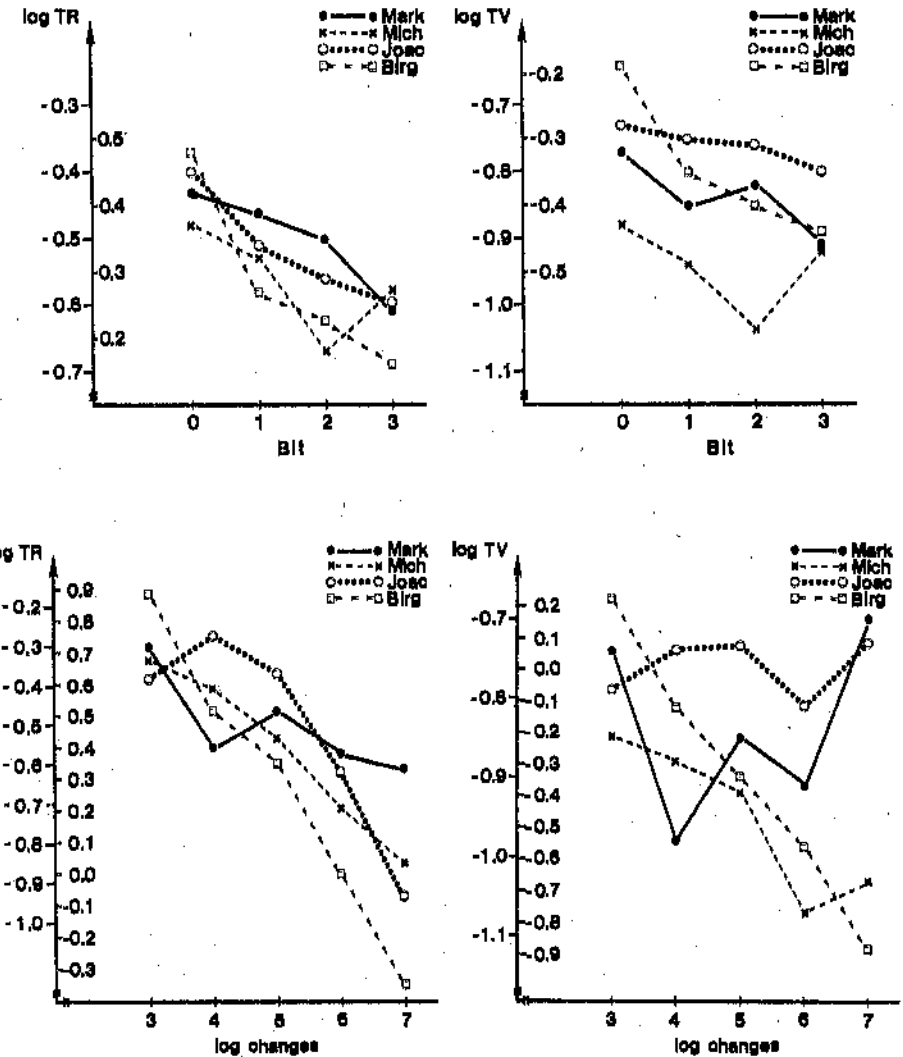


Abb. 1

Mittelwerte der logarithmierten Täuschungsbeträge für die subjektive Reproduktion (log TR) und die verbale Schätzung (log TV) unter den Stufen des Faktors „Effort“ (oberer Teil) sowie unter den Stufen des Faktors „Changes“ (unterer Teil). Die Changes werden als log dualis angegeben. Die links der Ordinate angegebene Skala gilt für die Pbn Mark, Mich und Joac, die rechts davon für Pb Birg.

Tabelle 3

Ergebnisse der F-Tests auf Haupteffekte „Changes“ (CH) und „Effort“ (PE) sowie deren Interaktion unter Einbezug der Kovariate „logarithmierte objektive Zeit“ (log OBZ), getrennt nach zwei abhängigen Variablen und nach den vier Pbn. Die Spalten H(CHM) und H(PEM) geben an, ob die Daten mit einem der beiden Modelle verträglich sind. Bei den eingeklammerten Werten hat der Test auf homogene Regressionen der AV auf die Kovariate innerhalb dieser UV Abweichungen nachgewiesen ($p \leq 0.10$).

AV*	F-Bruch des Effekts			Kovariat	R ²	H(CHM)	H(PEM)
	CH	PE	CH*PE	log OBZ			
				Birg			
log(TR)	0.92	(2.02*)	1.92*	31.60*	0.73	nein	(ja)
log(TV)	(1.40)	6.32*	1.31	80.03*	0.81	(nein)	ja
				Mark			
log(TR)	3.85*	10.64*	0.85	37.03*	0.44	nein	nein
log(TV)	0.32	5.05*	0.89	1.50	0.20	nein	ja
				Mich			
log(TR)	(1.06*)	4.29*	1.18	33.94*	0.45	(nein)	ja
log(TV)	(0.99)	1.81	0.73	10.16*	0.17	(nein)	nein
				Joac			
log(TR)	0.31	(9.74*)	1.28	11.60*	0.56	nein	(ja)
log(TV)	0.04	0.74	1.08	1.44	0.11	nein	nein

* log TR = logarithmierter Täuschungsbetrag bei Reproduktion; log TV = logarithmierter Täuschungsbetrag bei Verbalschätzung.

* $p \leq 0.10$ bei $df_{CH} = 3$, $df_{PE} = 4$, $df_{CH*PE} = 12$ und $df_{Error} = 139$.

zwei Ausnahmen erzielt die Kovariate bedeutsame Effekte; dies besagt, daß die Täuschungsbeträge log TR und log TV stark von log OBZ abhängig sind. Da die Sortierzeit in stärkerem Maß von „Changes“ als von „Effort“ beeinflusst wird, verschwinden unter Einbezug der Kovariate die in Tabelle 1 aufgezeigten Haupteffekte „Changes“ fast vollständig. Dennoch reichen diese Daten nicht aus, um von einer Bewährung etwa des Effort-Modells sprechen zu können. Vielmehr müßte man sagen, daß neben den direkten Treatment-Wirkungen auf die Täuschungsbeträge auch noch indirekte Wirkungen, vermittelt über die objektive Sortierzeit, vorliegen.

Insgesamt ist die resultierende Lage wie folgt zu charakterisieren: Die Hypothesen jeweils einseitig wirkender Faktoren (Effort oder Changes) als einander ausschließender Varianzquellen müssen aufgegeben werden zugunsten eines Modells, in dem beide Einflußfaktoren *additiv* zusammen-

wirken (der Interaktionstest erreicht nur bei Pb Birg bei log TR knapp die Signifikanzgrenze). Dieses additive Modell gilt in erster Linie für die nach der Methode der subjektiven Reproduktion erhobenen Daten. Das Urteilsverfahren der verbalen Schätzung ist bei zwei von vier Pbn einem Effekt der beiden Faktoren gegenüber insensitiv. Generell gilt, daß neben den direkten Auswirkungen der Faktoren auch noch indirekte Effekte vorliegen, die durch die variierende objektive Sortierzeit vermittelt sind.

Explorative Analysen

Die eben berichteten Befunde entstammen einem Versuchsplan, der gegenüber Störungen der internen Validität in besonderem Maße anfällig ist. Daher soll auf zwei derartige Störquellen, serielle Abhängigkeit und Meßwiederholungseffekte, näher eingegangen werden. Weiterhin soll geklärt werden, ob die Daten mit einer Potenzfunktion vereinbar sind. Schließlich werden die jeweiligen Informationsverarbeitungsdaten untersucht.

Serielle Abhängigkeit. Neben dem durch die unterschiedlichen Sortierbedingungen realisierten Kontext gibt es einen zeitlichen Kontext, in den das Urteil des Pb eingebettet ist: Es kann nicht ausgeschlossen werden, daß z. B. die gerade vorangegangene Versuchsbedingung noch Nachwirkungen auf die aktuelle Urteilssituation ausübt. Um einen derartigen Effekt serieller Abhängigkeit aufzufindig zu machen, wurden die Daten der Pbn in der Reihenfolge ihrer Erhebung zeitversetzt autokorreliert (die Berechnung erfolgte mit dem Modul „SERIES“ aus dem SYSTAT-Paket von Wilkinson, 1986). Die resultierenden partiellen Autokorrelationen für eine Zeitverschiebung um ein bis vier Takte zeigt Tabelle 4 sowohl für die Rohdaten als auch für die zwei Täuschungsmaße.

Die Ergebnisse dieser Analysen sprechen dafür, daß bei den Täuschungsmaßen bis auf eine Ausnahme (log TR bei Pb Mich) das Ausmaß der vorangegangenen Täuschung auf die aktuelle Täuschung Einfluß nimmt. Die vereinzelt auftretenden höheren Partialkorrelationen bei Verzögerungen um zwei bis vier Takte scheinen eher unsystematisch (von Pb Birg abgesehen, die recht konsistent auch noch lag-4-Einflüsse zeigt — hierfür kann keine Erklärung gegeben werden). Damit gehen im Schnitt ca. 10% der Varianz der Täuschungsbeträge zu Lasten serieller Abhängigkeit. Überraschend ist der Befund, daß die Rohdaten nicht unbedingt autokorreliert sein müssen: Bei Pb Joac sind sowohl reproduzierte Zeit als auch verbale Schätzung unabhängig vom jeweils vorangegangenen Urteil — dies gilt aber nicht für die Täuschungsbeträge!

Meßwiederholungseffekte. Von den eben beschriebenen Effekten aufeinanderfolgender Messungen sind die eigentlichen Meßwiederholungseffekte

Tabelle 4

Partielle Autokorrelationen für Lag 1 bis 4 von fünf abhängigen Variablen sowie Mittelwert (M) und Streuung (SD) der jeweiligen Zeitreihe, getrennt nach den vier Pbn.

AV ^a	M	SD	Lag			
			1	2	3	4
Birg						
OZ	88.49	77.42	.07	-.10	-.06	.23 ^b
RZ	88.15	57.49	.14	-.06	.06	.08
VS	47.12	28.46	.11	-.09	.00	.24 ^b
log(TR)	0.28	0.56	.21 ^b	.10	.06	.22 ^b
log(TV)	-0.34	0.46	.17 ^b	-.01	.07	.23 ^b
Mark						
OZ	81.51	72.37	.21 ^b	.05	.07	-.01
RZ	47.12	40.66	.15	.00	.03	-.03
VS	38.91	42.38	.24 ^b	.18 ^b	.07	.01
log(TR)	-0.50	0.29	.30 ^b	.15	.17 ^b	.23 ^b
log(TV)	-0.84	0.32	.34 ^b	.24 ^b	.25 ^b	.12
Mich						
OZ	87.17	76.13	.07	-.10	-.06	.23
RZ	44.39	35.48	.41 ^b	.00	.16 ^b	-.02
VS	33.89	31.76	.35 ^b	.04	.09	-.04
log(TR)	-0.57	0.40	.04	.01	.12	.06
log(TV)	-0.94	0.39	.21 ^b	.02	.04	-.04
Joaac						
OZ	80.74	70.82	.03	.13	-.14	.00
RZ	41.81	29.60	.02	.13	.00	.01
VS	39.19	36.05	.07	.11	-.17 ^b	.04
log(TR)	-0.52	0.35	.26 ^b	.15	-.01	.11
log(TV)	-0.76	0.31	.30 ^b	.07	.08	-.02

^a OZ = objektive Zeit in sec; RZ = reproduzierte Zeit in sec; VS = verbale Schätzung in sec; log TR = logarithmierter Täuschungsbetrag bei Reproduktion; log TV = logarithmierter Täuschungsbetrag bei Verbalschätzung.

^b Der markierte Wert liegt außerhalb des Bereichs $\pm 2 * \text{Standardfehler (SE)}$, wobei für alle Partialautokorrelationen gilt: SE = 0.079 bei jeweils N = 160 Datenpunkten.

zu unterscheiden, die sich auf die insgesamt achtfache Wiederholung jeder einzelnen Versuchsbedingung beziehen. Um etwaige Veränderungen im Urteils- bzw. Täuschungsverhalten sichtbar zu machen, zeigt Abbildung 2 für jeden Pb und für beide AVn die Verteilung der Daten über die acht Wiederholungen in Form von „box-and-whisker“-Plots nach Tukey (1977, Chap. 2C; Modul „GRAPH“ von Wilkinson, 1986).

Man erkennt an der Darstellung in Abbildung 2 recht gut, daß sich sowohl Verteilung als auch zentrale Tendenz bei einigen Pbn systematisch verändern (z. B. Pb Mich bei log TR; Pb Joac bei log TV). Trotz ausbleibender Rückmeldung verändert sich also das Täuschungsverhalten durch die bloße Tatsache wiederholter Messungen.

Potenzfunktion. Für die Beziehung von subjektiver (t_{sub}) und objektiver Zeit (t_{obj}) wird oftmals eine psychophysikalische Funktion im Sinne des Potenzgesetzes von Stevens postuliert. Eisler (1976) gibt nach Durchsicht von 111 einschlägigen Studien an, daß der Exponent b der Potenzfunktion

$$t_{\text{sub}} = a * \log t_{\text{obj}}^b \quad (1)$$

konsistent Werte zwischen 0.5 und 1 annimmt. Michon (1985, S. 40) dagegen meint, daß nur zwei verschiedene Exponenten, 0.5 für eine „impressionistische“ Skala bei objektiven Zeiten kleiner als 0.5 oder größer als 20 sec und 1 für eine „analytische“ Skala im Bereich von 0.5 bis 20 sec, anzunehmen seien und alle Zwischenwerte das Resultat von über Pbn aggregierten Daten bzw. gemischten Strategien der Pbn wären. Tabelle 5 zeigt die Parameter a und b , die bei individueller Anpassung von jeweils 160 Datenpunkten an das Stevens-Modell (Gleichung 1) eruiert wurden.

Die Exponenten b entstammen bei beiden AVn dem angegebenen Wertebereich. Die oben angeführte Vermutung Michons über die Existenz nur

Tabelle 5

Koeffizienten der in Gleichung 1 angegebenen Potenzfunktion für vier Pbn mit je 160 Datenpunkten und für zwei verschiedene AVn. R^2 gibt die durch die Modellfunktion aufgeklärte Varianz an.

Pb	Parameter		R^2
	a	b	
(a) reproduzierte Zeit			
Birg	2.162	0.532	0.706
Mark	-0.012	0.878	0.899
Mich	0.420	0.757	0.811
Joaac	0.398	0.770	0.870
(b) verbale Schätzung			
Birg	1.295	0.591	0.859
Mark	-0.901	1.016	0.897
Mich	-0.516	0.894	0.827
Joaac	-0.765	1.001	0.903

zweier Exponenten auf der Ebene von Individualdaten muß allerdings zurückgewiesen werden. Die durch das Modell der Gleichung (1) aufgeklärte Varianz läßt die Vermutung zu, daß noch andere Prädiktoren in Erwägung zu ziehen wären. Die Anpassungsgüte ist nicht überwältigend.

Naheliegender ist die Frage, inwiefern Konstante a und Exponent b der Potenzfunktion von der realisierten „Kontextbedingung“, d.h. den verschiedenen Effort-Stufen abhängig sind. Das um den Kontextindex c erweiterte Modell lautet in diesem Fall:

$$t_{\text{sub}(c)} = a_{(c)} * \log t_{\text{obj}}^{b(c)} \quad (2)$$

Getestet werden kann dieses in Gleichung 2 beschriebene Modell dadurch, daß man den durch das Kontext-Modell erzielten Gewinn an erklärter Varianz auf Signifikanz prüft. Bei log RZ, der subjektiv reproduzierten Zeit, erwies sich dieser Gewinn für Pb Birg und Pb Joac als bedeutsam, bei log VS dagegen, der verbalen Schätzung, in keinem der vier Fälle. Die genauen Angaben sind beim Autor erhältlich.

Für die mit der Reproduktionsmethode gewonnenen Daten ist somit eine kontextabhängige Potenzfunktion gemäß Gleichung 2 anzunehmen. Dabei gibt es solche Pbn, bei denen der Kontext gleich wirkt (die Parameter unterscheiden sich bei ihnen also nicht), und solche, bei denen die Parameter verschieden ausfallen. Das allgemeinere Modell der Gleichung 2 ist somit dem in Gleichung 1 formulierten Spezialfall vorzuziehen.

Informationsverarbeitungsrate. Hicks et al. (1976, S. 723) berichten von durchschnittlichen Informationsverarbeitungsrate in der Größenordnung von etwa 1 bit in der 1-Bit- und von 1.55 bit in der 2-Bit-Bedingung. Bei dem dort verwendeten 42-sec-Intervall entspricht dies 45.8 bzw. 34.8 durchschnittlich sortierten Karten. In der 0-Bit-Bedingung wurden dort sogar im Mittel 75.2 Karten bearbeitet. Das Verhältnis „bearbeitete Karten pro Sekunde“ (KAPS) liegt also bei Hicks et al. im Mittel bei 1.79, 1.09 bzw. 0.83 für die drei verschiedenen Bit-Bedingungen. Im vorliegenden Experiment wurde wesentlich langsamer gearbeitet. Dies mag an der spezifischen Bearbeitungsweise (Entnahme der Karten aus dem Kartenschlitten, Umdrehen, Einsortieren in entsprechendes Fach) liegen, die eine gewisse motorische Leistung verlangte. Varianzanalysen mit der AV „KAPS“ zeigen bei allen vier Pbn signifikante Haupteffekte der beiden Faktoren bei gleichzeitig fehlendem Interaktionseffekt. Daß die Bearbeitungsrate über die vier verschiedenen Bitbedingungen fast linear abfällt (Mittelwerte von KAPS bei jeweils $N = 40$ für Pb Birg: 0.68, 0.61, 0.52, 0.49; für Pb Mark: 0.74, 0.59, 0.54, 0.49; für Pb Mich: 0.65, 0.56, 0.50, 0.47; für Pb Joac: 0.69, 0.63, 0.57, 0.51), spricht für die Wirksamkeit dieser experimentellen Manipulation. Daß dagegen auch die Changes-Bedingungen auf KAPS substantielle Effekte zeigen (Mittelwerte von KAPS über alle Pbn hinweg in den

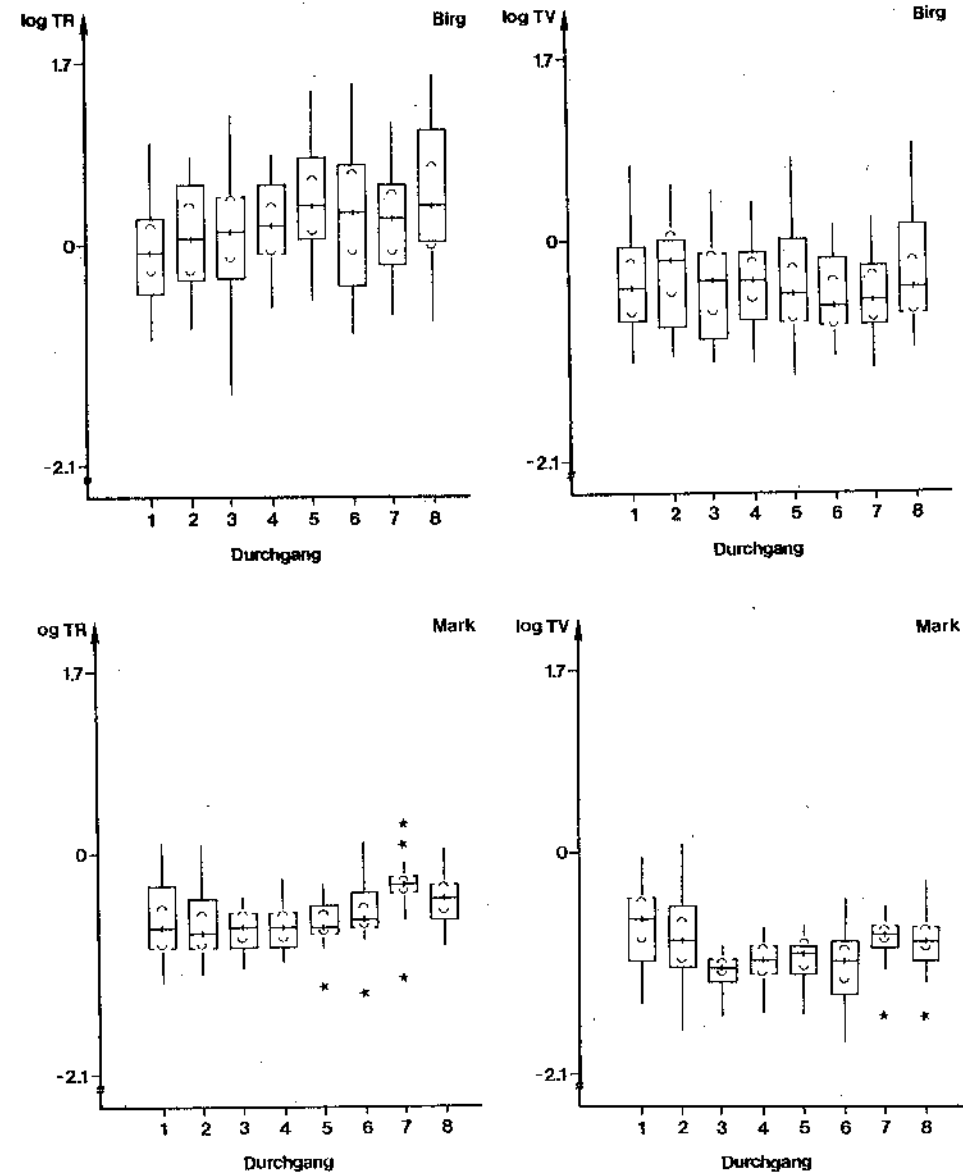


Abb. 2

Graphische Darstellung von Meßwiederholungseffekten in Form von „box-and-whisker“-Plots bei vier Pbn und zwei AVn. Die runden Klammern geben das 95%-Konfidenzintervall des Medians an, der in der Mitte jedes Kastens eingetragen ist. Die Kastenränder entsprechen dem ersten und dritten Quartil, Ausreißer sind mit einem Stern markiert.

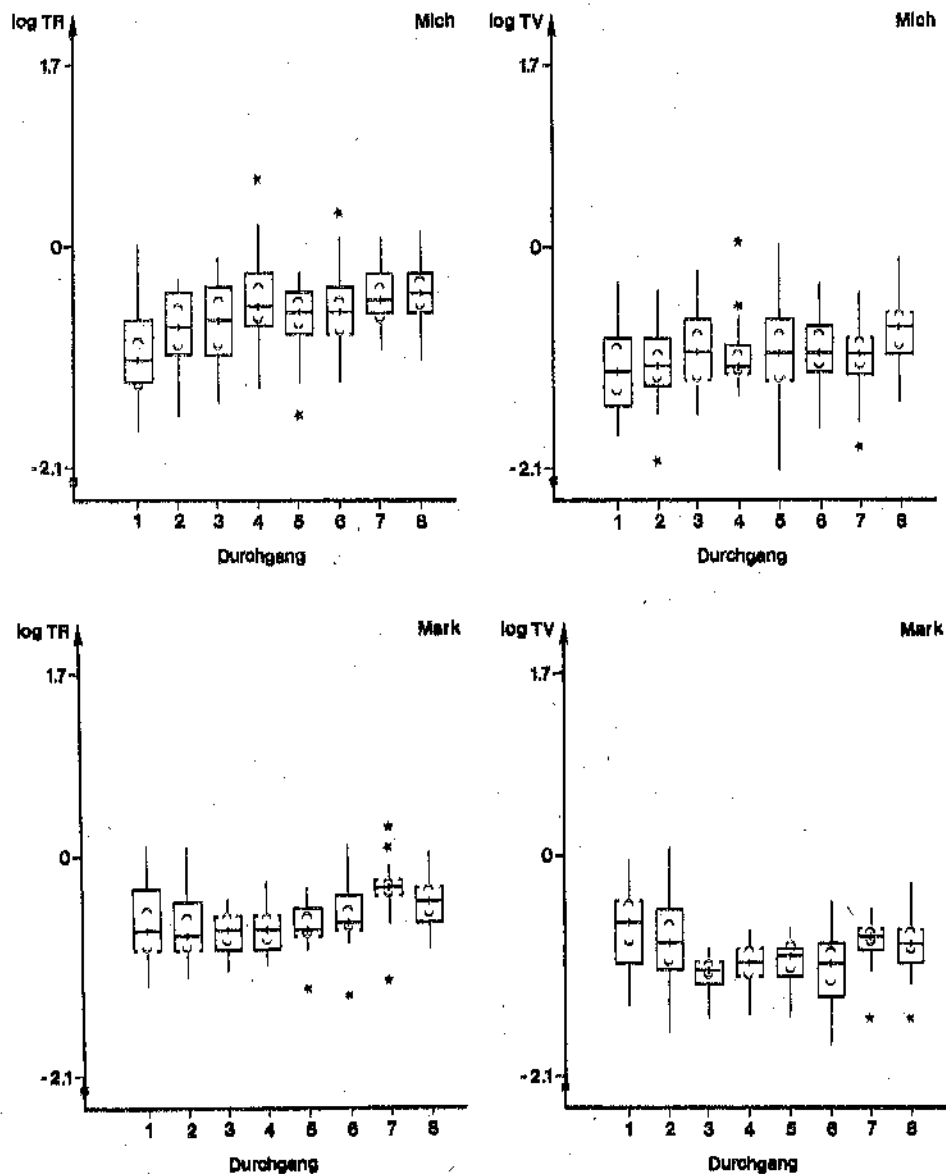


Abb. 2

fünf aufsteigend angegebenen Bedingungen: 0.52, 0.56, 0.61, 0.64, 0.64; jeweils $N = 128$), war nicht beabsichtigt. Die beschleunigte Bearbeitung bei wachsender Kartenzahl spricht für den Aufbau motorischer Routinen in der aktuellen Versuchssituation.

Ein derartiger Effekt ist im übrigen auch versuchsübergreifend nachzuweisen. Die KAPS-Rate über die acht Meßwiederholungen jeder Bedingung (Daten aller vier Pbn zusammengenommen) zeigt ebenfalls einen bedeutsamen F-Wert von 2.05 ($df = 7, 631$). Die entsprechenden Mittelwerte, basierend auf jeweils 80 Daten, wachsen wie folgt an: 0.55, 0.56, 0.56, 0.64, 0.61, 0.60, 0.61, 0.62. Die beim Kartensortieren erforderlichen motorischen Fertigkeiten verbessern sich danach wohl im Verlauf des Experiments.

Diskussion

Die eben berichteten Ergebnisse machen deutlich, daß Theorien zur Zeitschätzung noch keineswegs befriedigend das Zustandekommen von Zeittäuschungen, wie sie in der vorliegenden Arbeit untersucht wurden, erklären können. Die doch recht einfachen Haupteffekt-Modelle erweisen sich bei isolierter Betrachtung als nicht mit den Daten verträglich. Faßt man sie jedoch zu einem Modell additiver Wirkung mit direkten und indirekten Effekten zusammen, verbessert sich die Situation.

Es bleibt die Frage, inwiefern ein methodisches Vorgehen wie das hier gewählte — $N = 1$ -Design mit Meßwiederholung — trotz problematischer Aspekte (serielle Abhängigkeit, Lerneffekte, usw.) nicht doch den günstigsten Weg zur Prüfung solcher Theorien darstellt, die auf Individuenebene Gültigkeit beanspruchen. Gerade die interindividuellen Unterschiede (etwa bezüglich der Kontextsensitivität von Parametern der Potenzfunktion) zeigen deutlich die Grenzen aggregierter Datenanalysen auf. Die intraindividuelle Stabilität ihrerseits sorgt dafür, daß dennoch zuverlässige Aussagen über den Phänomenbereich möglich sind.

Summary

Theories concerning time estimation vary with respect to the factors considered to be influential in the realization of time illusions. Cognitive models consider as determinants the quantity of the changes having occurred or the extent of the processing effort invested within a certain period of time. In a conclusive experiment as well as in three replication studies (respectively $N = 1$), it was shown that neither changes nor effort are the sole source of influence. In contrast to other works, the present analysis

shows how a confounding problem can eventually be solved which up until now has hindered the independent variation of the time estimation determinants under examination. The results also point to great interindividual differences together with a high intraindividual consistency. This is evidenced by the problematic nature of the aggregation of data over individuals.

Literatur

- Aschoff, J. (1985). Zur Zeitwahrnehmung während langdauernder Isolation. *Human Neurobiology*, 4, 41—52.
- Bindra, D. & Waksberg, H. (1956). Methods and terminology in studies of time estimation. *Psychological Bulletin*, 53, 55—59.
- Block, R. A. (1974). Memory and the experience of duration in retrospect. *Memory and Cognition*, 2, 153—160.
- Block, R. A. (1985). Contextual coding in memory: Studies of remembered duration. In J. A. Michon & J. L. Jackson (Hrsg.), *Time, mind, and behavior* (S. 169—178). Heidelberg: Springer.
- Block, R. A. & Reed, M. A. (1978). Remembered duration: Evidence for a contextual-change hypothesis. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 656—665.
- Bobko, D. J., Schiffman, H.R., Castino, R. J. & Chiapetta, W. (1977). Contextual effects in duration experience. *American Journal of Psychology*, 90, 577—586.
- Borg, I. & Galinat, W. (1985). Der Einfluß von Merkmalen der Situation auf das Erleben ihrer Dauer. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 32, 353—369.
- Bredenkamp, J. (1980). *Theorie und Planung psychologischer Experimente*. Darmstadt: Steinkopff.
- Bredenkamp, J. (1984). Theoretische und experimentelle Analyse dreier Wahrnehmungstauschungen. *Zeitschrift für Psychologie*, 192, 47—61.
- Brown, S. W. (1985). Time perception and attention: The effects of prospective versus retrospective paradigms and task demands on perceived duration. *Perception & Psychophysics*, 38, 115—124.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.
- Eisler, H. (1976). Experiments on subjective duration 1868—1975: A collection of power function exponents. *Psychological Bulletin*, 83, 1154—1171.
- Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, 35, 1—36.
- Fraisse, P. (1985). *Psychologie der Zeit*. Basel: Reinhardt.
- Guay, M. & Bourgeois, J. (1981). Short term retention of temporal information. *Perceptual & Motor Skills*, 52, 719—726.
- Hager, W. (1987). Grundlagen einer Versuchsplanung zur Prüfung empirischer Hypothesen der Psychologie. In G. Lüer (Hrsg.), *Allgemeine Experimentelle Psychologie. Eine Einführung in die methodischen Grundlagen mit praktischen Übungen für das Experimentelle Praktikum* (S. 43—264). Stuttgart: Gustav Fischer.
- Hicks, R. E., Miller, G. W. & Kinsbourne, M. (1976). Prospective and retrospective judgments of time as a function of amount of information processed. *American Journal of Psychology*, 89, 719—730.
- Hornstein, A. D. & Rotter, G. S. (1969). Research methodology in temporal perception. *Journal of Experimental Psychology*, 79, 561—564.

- James, W. (1890). *The Principles of Psychology*. New York: Holt.
- Leichsenring, F. (1987). Einzelfallanalyse und Strenge der Prüfung. *Diagnostica*, 33, 93—109.
- Michon, J. A. (1985). The complex time experimenter. In J. A. Michon & J. L. Jackson (Hrsg.), *Time, mind, and behavior* (S. 20—52). Heidelberg: Springer.
- Miller, G. W., Hicks, R. & Willette, M. V. (1978). Effects on concurrent verbal rehearsal and temporal set upon judgements of temporal duration. *Acta Psychologica*, 42, 173—179.
- Mulligan, R. M. & Schiffman, H. R. (1979). Temporal experience as a function of organisation in memory. *Bulletin of the Psychonomic Society*, 14, 417—420.
- Ornstein, R. E. (1969). *On the experience of time*. Baltimore: Penguin Books.
- Ornstein, R. E. (1974). *Psychologie des Bewußtseins*. Köln: Kiepenheuer.
- Poynter, W. D. & Homa, D. (1983). Duration judgment and the experience of change. *Perception & Psychophysics*, 33, 548—560.
- Thomas, E. A. C. & Weaver, W. B. (1975). Cognitive processing and time perception. *Perception & Psychophysics*, 17, 363—367.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, Mass.: Addison-Wesley.
- Vroon, P. A. (1970). Effects of presented and processed information on duration experience. *Acta Psychologica*, 34, 115—121.
- Westermann, R. & Hager, W. (1986). Error probabilities in educational and psychological research. *Journal of Educational Statistics*, 11, 117—146.
- Wilkinson, L. (1986). *SYSTAT: The system for statistics*. Evanston, Ill.: SYSTAT Inc.

Anschrift des Verfassers: Dr. Joachim Funke, Psychologisches Institut, Römerstr. 164, D-5300 Bonn 1.