

- Spence, S. A. (2008). Playing the devil's advocate: The case against fMRI lie detection. *Legal and Criminological Psychology*, 13, 11–25.
- Stahl, J. & Rammsayer, T. (2004). Differences in the transmission of sensory input into motor output between introverts and extraverts: Behavioral and psychophysiological analyses. *Brain & Cognition*, 56, 293–303.
- Stelmack, R. M. & Houlihan, M. (1995). Event-related potentials, personality and intelligence: Concepts, issues and evidence. In D. H. Saklofske & M. Zaidner (Eds.), *International handbook of personality and intelligence* (pp. 349–365). New York: Plenum.
- Stelmack, R. M. & Rammsayer, T. (2008). Psychophysiological and biochemical correlates of personality. In G. J. Boyle, G. Matthews & D. H. Saklofske (Eds.), *Handbook of personality theory and assessment* (Vol. 1., pp. 33–55). London: Sage.
- Stemmler, G. (2003). Methodological considerations in the psychophysiological study of emotion. In R. J. Davidson, H. H. Goldsmith & K. R. Scherer (Eds.), *Handbook of affective science* (pp. 225–255). New York: Oxford University Press.
- Stemmler, G. (2004). Physiological processes during emotion. In P. Philippot & R. S. Feldman (Eds.), *The regulation of emotion* (pp. 33–70). London: Erlbaum.
- Surwillo, W. W. & Quilter, R. E. (1965). The relation of frequency of spontaneous skin potential responses to vigilance and to age. *Psychophysiology*, 1, 272–276.
- Thatcher, R. W., North, D. & Biver, C. (2005). EEG and intelligence: Relations between EEG coherence, EEG phase delay and power. *Clinical Neurophysiology*, 116, 2129–2141.
- Traxel, W., Gundlach, H. & Zschuppe, U. (1986). Zur Geschichte der apparativen Hilfsmittel der Psychologie. In R. Brickenkamp (Hrsg.), *Handbuch apparativer Verfahren in der Psychologie* (S. 1–22). Göttingen: Hogrefe.
- Turner, S. M., DeMers, S. T., Fox, H. R. & Reed, G. M. (2001). APA's guidelines for test user qualifications. *American Psychologist*, 56, 1099–1113.
- Ulrich, R. & Giray, M. (1989). Time resolution of clocks: Effects on reaction time measurement. Good news for bad clocks. *British Journal of Mathematical and Statistical Psychology*, 42, 1–12.
- Vogt, J., Hagemann, T. & Kastner, M. (2006). The impact of workload on heart rate and blood pressure in en-route and tower air traffic control. *Journal of Psychophysiology*, 20, 297–314.
- Youn, T., Lyoo, I. K., Kim, J.-J., Park, H.-J., Ha, K.-S., Lee, D. S. et al. (2002). Relationship between personality trait and regional cerebral glucose metabolism assessed with positron emission tomography. *Biological Psychology*, 60, 109–120.
- Zakzanis, K. K., Graham, S. J. & Campbell, Z. (2003). A meta-analysis of structural and functional brain imaging in dementia of the Alzheimer's type: A neuroimaging profile. *Neuropsychology Review*, 13, 1–18.
- Zipes, D. P. & Jalife, J. (2000). *Cardiac electrophysiology: From cell to bedside*. Philadelphia, PA: Saunders.

## 12. Kapitel

## Ambulantes Assessment

Bernd Reuschenbach und Joachim Funke

## 1 Einleitung

Ein Grundmotiv differenzialpsychologischer Forschung ist die Aufdeckung individueller Verhaltens- und Erlebensaspekte, aus der sich beispielsweise Empfehlungen für Interventionen ergeben. Bevorzugt werden dazu Befragungen mittels strukturierter und/oder standardisierter Fragebögen oder Interviews genutzt. Die Idee, aus dem Ankrenzverhalten in Fragebögen Rückschlüsse auf das zurückliegende oder zukünftige Verhalten ziehen zu können, ist eine wesentliche Annahme der Diagnostik, die durch vielfältige Korrelationsstudien eine empirische Unterstützung erfährt (Funder, 2001). Mit der Nutzung von Fragebögen im Labor werden situationale Einflüsse reduziert. An die Stelle einer unmittelbaren Erfassung des Verhaltens in der natürlichen Umgebung tritt die verzögerte Erfassung mentaler Repräsentationen des Verhaltens. Diese verzögerte Erfassung ist mit typischen Verzerrungen (z. B. „recall bias“) verbunden und reduziert die Validität der Ergebnisse und letztlich auch die Vitalität des psychologischen Tuns, das sich vom realen Leben und menschlichen Verhalten zunehmend verabschiedet, um sich der Auswertung von Fragebögen hinzugeben. Pointiert bringt dies ein Beitrag von Baumeister, Vohs und Funder (2007) auf den Punkt, die nach einer Analyse der Beiträge im *Journal of Personality and Social Psychology* feststellen: „Direct observation of meaningful behavior is apparently passé“ (S. 397). Wenn die in diesem Journal enthaltenen Beiträge ein Spiegelbild des psychologischen Handelns sind, „then human behavior is almost always performed in a seated position, usually seated in front of a computer. Finger movements, as in keystrokes and pencil marks, constitute the vast majority of human action“ (S. 397).

Die darin implizit enthaltene und im weiteren Teil ihres Beitrags sehr explizit vorgetragene Forderung nach einer Rückbesinnung auf das eigentliche Verhalten

und das unmittelbare Erleben in der gewohnten Umgebung wird im Ansatz des ambulanten Assessment aufgegriffen. Dieser Ansatz ist nicht neu, erlebt aber seit den 1990er Jahren einen regelrechten Boom und es ist davon auszugehen, dass diese Messstrategie die psychologische Diagnostik der nächsten Jahre maßgeblich beeinflussen wird. In den folgenden Abschnitten sollen daher zunächst die wichtigsten Kennzeichen des ambulanten Assessments, die methodischen Schwierigkeiten und einige dominierende Einsatzmöglichkeiten genannt werden. Darüber hinaus werden auch innovative technische Instrumentarien vorgestellt werden, die in Zukunft für das ambulante Messen an Bedeutung gewinnen werden. Das Kapitel schließt mit einer Abwägung der Vor- und Nachteile des ambulanten Assessments. Dabei werden insbesondere die ethischen Aspekte, die sich durch eine unaufdringliche Aufzeichnung von Daten ergeben, thematisiert.

## 2 Ambulantes Assessment: Definitionen und Abgrenzungen

„Ambulantes Assessment“, auch „ambulatory assessment“ oder „ambulatory monitoring“ genannt, ist eine Messstrategie, kein spezielles Messverfahren. „Ambulant“ verweist auf die Erhebung im Feld, die meist mehrfach und außerhalb der Reichweite eines Testleiters stattfindet. Im weiteren Sinne kann unter „ambulatory assessment“ schon die Nutzung von Papier-Bleistift-Fragebögen oder Tagebüchern außerhalb des Labors (z. B. in der häuslichen Umgebung) verstanden werden. Synonym wird daher auch gelegentlich von Tagebuchverfahren oder „diary methods“ gesprochen (vgl. Hoppmann & Riediger, 2009). Im engeren Sinne ist hiermit jedoch zusätzlich die Verwendung von Geräten (z. B. von PDA-Geräten<sup>1</sup> oder tragbaren bzw. funkbasierten physiologischen Messgeräten) zur Datenerhebung gemeint.

Die Idee, psychologische und auch physiologische Daten außerhalb des Labors in der natürlichen Umgebung – bei Kurt Lewin als „Feld“, von Cattell als „in situ“ bezeichnet – zu sammeln, ist nicht neu. Frühe Arbeiten stammen aus der Ethnografie, bei der die Entstehung sozialer Muster über die Zeit hinweg untersucht wurde (Malinowski, 1935). Ökopsychologische Studien von Barker und Wright (1951, 1955) sind wichtige Pionierarbeiten, die für die Entwicklung des ambulanten Monitorings oder Assessments wegweisend waren. Mit der von ihnen formulierten „ecological psychology“ wird ein theoriefreies und naives Sammeln von Verhaltensdaten in der natürlichen Umgebung gefordert, um die Komplexität des Handelns zu erschließen. Dazu gründeten sie 1947 die Forschungsstation „Midwest Psychological Field Station“ in Oskaloosa in Kansas, um ein ganzes Jahr lang das Alltagsleben der 725 Einwohner zu erfassen (Lichtenberg, Eitmann & Goldmann, 2003). Zunächst stand der Verhaltens-

1 PDA: Personal Digital Assistant

strom („stream of behavior“) von Kindern im Mittelpunkt, der von morgens bis abends mitsamt den entsprechenden zeitlichen Mustern und situativen Kontexten beobachtet wurde. Es zeigte sich, dass Personen typische Verhaltensmuster aufweisen, die durch den sozialen und materiellen Kontext bestimmt waren (sog. „behavior-settings“): „The behavior of different children within the same region was often more similar than the behavior of any of them in different regions“ (Barker, 1968, S. 203). Aus den Erkenntnissen schlussfolgern die Autoren, dass nur eine zeitnahe Erfassung in der natürlichen Umgebung die relevanten Situations-Verhaltens-Verbindungen aufdecken kann.

In Deutschland wurden erste psychologische Studien zum ambulanten Assessment schon in den 1980er Jahren durchgeführt (z. B. Pawlik & Buse, 1982). In den USA bekam die Messstrategie insbesondere durch die von Larson und Csikszentmihalyi (1983) beschriebene *Experience-Sampling-Methode* (ESM) einen wesentlichen Auftrieb. „Experience Sampling Method (ESM), [is] a research procedure that consists of asking individuals to provide systematic self-reports at random occasions during the waking hours of a normal week“ (S. 41). Die Erhebungen fanden anfänglich mit klassischen Papier-Bleistift-Bögen statt; lediglich der Hinweis an die Testpersonen zur Erfassung der Dateneingabe, das sogenannte Prompting, erfolgte mittels eines Gerätes oder durch einen Zeitgeber. Erst später erfolgte dann auch die Dateneingabe elektronisch. Als wesentliche Vorteile dieser „naturalistischen“ Erfassung von Verhaltens- und Erlebnisaspekten wird von den Autoren die fehlende Verzerrung durch die retrospektive Rekonstruktion der Sachverhalte genannt (Käppler & Rieder, 2001). Daten werden zeitnah und in der natürlichen Umgebung erhoben.

Es waren vor allem die zunehmende Miniaturisierung elektronischer Geräte, die nachlassenden Kosten für elektronische Geräte, die Möglichkeit zur digitalen Datenspeicherung und die Weiterentwicklung der drahtlosen Übertragungstechniken (WLAN<sup>2</sup>, Bluetooth etc.), die die Messung im Felde vorantrieben. Parallel zu diesen technischen Veränderungen gab es auch in der Statistik Weiterentwicklungen, die es ermöglichten, durch geeignete Varianzdekompositionen den Einfluss von Trait, State, Situation und entsprechende Interaktionen im Hinblick auf die erhobenen Testwerte adäquat zu bestimmen.

Für den deutschsprachigen Raum hat die Forschungsgruppe um Fahrenberg und Myrtek (Fahrenberg & Myrtek, 1996, 2001) wichtige Pionierarbeit geleistet.

Mit der *Society for Ambulatory Assessment* (SAA) wurde im Jahre 2008 die erste Fachgesellschaft gegründet, die sich zum Ziel gesetzt hat, die Forschungsmethode weiter zu fördern und die Vernetzung der Anwender voranzutreiben.

2 WLAN: Wireless Local Area Network

Die Internetseite [www.ambulatory-assessment.org](http://www.ambulatory-assessment.org) enthält umfangreiche Zusammenstellungen aktueller Forschungsarbeiten, Software- und Hardwareangebote sowie eine Expertendatenbank.

Fahrenberg, Myrtek, Pawlik und Perrez (2007) definieren den Begriff des ambulanten Assessments als „die Verwendung spezieller feldtauglicher, heute meist elektronischer Geräte und computer-unterstützter Erhebungsmethoden, um Selbstberichtsdaten, Verhaltensbeobachtungsdaten, psychometrische Verhaltensdaten, physiologische Messdaten, psychometrische Verhaltensmaße, physiologische Messwerte sowie situative und Setting-Bedingungen im Alltag der Untersuchten zu erfassen“ (S. 13). Die Begriffe *ambulatory assessment* und *ambulatory monitoring* werden häufig synonym gebraucht. Es wird jedoch deutlich, dass „monitoring“ eher die kontinuierliche Messung zum Beispiel von physiologischen Daten bezeichnet, während sich das „assessment“ vornehmlich auf die Eingabe von Daten in Tagebüchern oder digitalen Fragebögen durch die Testperson bezieht. Für den deutschsprachigen Raum empfehlen Fahrenberg et al. (2007) den Begriff des ambulanten Assessments. Weil Assessment jede Form der datengewinnenden Erhebungen bezeichnet (Reuschenbach & Mahler, in Druck), ist damit auch das automatisierte Monitoring eingeschlossen. Verwandte Begriffe, die jedoch auf die spezielle Erhebungsmethode, die Stichprobenbildung oder die Häufigkeit der Datenerhebung Bezug nehmen, sind: „diary method, event sampling method, beeper studies, real-time reporting techniques, time situated methods, intensive repeated measures methods, or ecological momentary assessment“ (Hoppmann & Riediger, 2009).

Ambulantes Assessment kann durch das Tragen entsprechender Messgeräte (häufig Pocket-PCs/PDAs) oder durch die telemetrische Übertragung der Daten stattfinden. Der Begriff *Telemetrie* bezieht sich zum einen auf die Übertragung von Daten mittels drahtlosen Sendevorrichtungen. Zum anderen wird aber auch die Eingabe von selbst ermittelten Daten (z. B. Blutzuckerwerte) oder Stimmungseinschätzungen über das Telefon als Telemetrie bezeichnet. Der Begriff findet sich häufiger im medizinischen Kontext (Telemedizin), bei dem die Messung meist entsprechende therapeutische Konsequenzen nach sich zieht. „Telemedicine is the use of information and telecommunication technologies to provide and support healthcare when distance separates the participants“ (Field, 1997, S. 28). Eine neuere Definition sieht Telemedizin als „any medical activity involving an element of distance and use of a telecommunications strategy“ (Farmner, Gibson, Tarassenko & Neil, 2005). Zu den „activities“ zählen in diesem erweiterten Begriffsverständnis Telemonitoring, Teleedukation, Teleconsulting und Telecare.

Im angloamerikanischen Raum wurde von Stone und Shiffman (1994) der Begriff *ecological momentary assessment* (EMA) eingeführt. Hufford, Shiffman,

Paty und Stone (2001) bezeichnen EMA als „konzeptuelle Strategie“, die durch eine Veränderung von „retrospective to momentary data collection“ gekennzeichnet ist. Folgende drei Aspekte kennzeichnen diesen Ansatz (vgl. Stone, Shiffman & DeVries, 1999):

1. Messung und Datenerfassung in der Umwelt, in der das Verhalten vorkommt. Dies sichert die ökologische Validität.
2. Echtzeitdatenmessung, also eine „momentary“ oder „near-immediate“ Erfassung der Daten: „EMA differs from other self-report methods in its focus on subjects' current immediate state. That is, in an EMA study of mood, subjects are not asked about how they feel in general or how they felt yesterday, but rather how they feel right now“ (Shiffman & Stone, 1998, S. 118). Mögliche Verzerrungseffekte, die sich durch die Retrospektion ergeben, werden dadurch minimiert.
3. Einsatz einer bestimmten Samplingstrategie zur Datenerfassung. Es können fünf verschiedene „recording paradigms“ verwendet werden: Event-, Signal- oder intervallkontingentes Sampling, weiterhin ein kontinuierliches Sampling und ein kontextsensitives Sampling (ausführlicher dazu vgl. Abschnitt 5.1).

Beim Vergleich der Definitionen von *ambulatory assessment* und EMA zeigt sich, dass beim EMA genauere Angaben zu den definierenden Merkmalen gemacht werden. In der praktischen Umsetzung sind beide Ansätze kaum zu unterscheiden. Es ist selbstverständlich, dass auch bei Forschungsansätzen, die unter der Überschrift *ambulantes Assessment* auftauchen, besondere Samplingstrategien eingehalten werden und eine zeitnahe Erfassung stattfindet. Zusammen mit den unterschiedlichen Datenquellen (Suls & Martin, 1993) lassen sich die Variationen des ambulanten Assessments auf zwei Dimensionen – Berichtsarten und Sampling-Strategien – aufspannen (vgl. Abb. 1).

	self-report	surveillance-report	automated-report
Event-Sampling			
Signal-Sampling			
Intervall-Sampling			
Kontinuierliches Sampling			
Kontext-sensitives Sampling			

**Abbildung 1:** Varianten des ambulanten Assessments, die sich aus der Kombination von fünf Sampling-Strategien und drei Berichtsarten ergeben

Hinsichtlich der „report source“ werden drei Arten unterschieden: (1) Zum „self-report“ zählt z. B. das Führen eines elektronischen Tagebuchs, (2) beim „surveillance-report“ übernehmen andere Personen die Messung und Dokumentation der Daten, (3) beim „automated report“ werden automatisiert Daten aufgenommen. Dies erfolgt beispielsweise bei physiologischen Messungen. Heger (1990) unterscheidet im Hinblick auf die Partizipation der Testpersonen *aktive Messungen*, bei der die Person an der Erhebung selbst mitwirkt, z. B. beim „self-monitoring“, *passive Messungen*, z. B. bei physiologischen Messungen, und *reaktive Messungen*, bei denen durch ein Signal oder die eigene Symptomwahrnehmung die Person zur Durchführung von Messungen animiert wird.

Eine weitere Einteilung ergibt sich anhand der psychologischen Konstrukte, die mittels ambulantem Assessment gemessen werden können. Tabelle 1 vermittelt einen Eindruck von den Möglichkeiten.

**Tabelle 1:**

Psychische Konstrukte, für die Strategien des ambulanten Assessments verwendet werden können

Konstrukt	Methode/Beispielstudie
Motorik und Körperlage („posture“)	<i>Akzelometrie:</i> – Bedeutsame Korrelate und Kontrollvariablen für andere physiologische Messung (Bussmann, van den Berg-Emons & Stam, 2001; Jain, Martens, Mutz, Weiß & Stephan, 1996) – Messung der Beweglichkeit in der Rehabilitationspsychologie, z. B. mittels „activity monitor“ (J. B. Bussmann et al., 2001) – Messung des Energieverbrauchs mittels „CALTRAC“ (Pambianco, Wing & Robertson, 1990) – Messung spezieller Körperbewegung, wie z. B. Tics mittels „Wrist-Actigraphy“ (Tulen, Vokers, Stronks, Cavelaars & Groenevela, 2001) – Messung des Aktivitätsniveaus (Bouten, Westerterp & Verduin, 1994; Fox, Stathi, McKenna & Davis, 2007) – Messung der körperlichen Aktivitäten mittels Pedometer (Bennett, Wolin, Puleo & Emmons, 2006)
Angst, Erregung	– Messung von Blutdruck (Höhe, Reaktivität), Herzfrequenz, EDA mittels mobilen Geräten (Myrtek, 2004)
Schmerz	– Elektronisches Schmerztagebuch zur unmittelbaren Erfassung des Schmerzes und des Schmerzverlaufs: „pain diaries“ (Jamison et al., 2004; Stone & Broderick, 2007)
Wachheit	– <i>Electronic Pilot-Activity Monitor</i> (EPAM) zur Messung der Wachheit von Piloten (Cabon et al., 2003) – Brain-Computer-Interface zur mobilen Messung von EEG-Strömen zur Messung der Wachheit (Lin et al., 2008)

**Tabelle 1** (Fortsetzung):

Psychische Konstrukte, für die Strategien des ambulanten Assessments verwendet werden können

Konstrukt	Methode/Beispielstudie
Schlaf	– Erfassung der circadianen Periodik (Speichelmelatonin, Körpertemperatur) in Abhängigkeit von der Lichtexposition, mittels tragbarer Lichtdosimeter (Goulet, Mongrain, Desrosiers, Paquet & Dumont, 2007) – Erfassung von Schlafstörungen mittels mobilem und telemetrisch übertragenen EEG-Vorrichtungen (z. B. Schomer, 2006)
Fatigue	– Messung von Fatigue mittels mobilem EMA-Monitor bei Chemotherapiepatienten (Dimsdale, Ancoli-Israel, Ayalon, Elmsore & Green, 2007) – Messung von Fatigue beim chronischen Fatigue-Syndrom (Friedberg & Sohl, 2009)
Stimmung/ Emotionen	– <i>Messung im Felde mittels PDA</i> (Hank, Schwenkmezger & Schumann, 2001; Jamner, Shapiro & Alberts, 1998) – <i>Messung von Emotionen mittels PDA</i> , z. B. vor, während und nach sportlichen Wettkämpfen (Cerin & Barnett, 2006) – <i>Messung von Basisemotionen</i> mittels PDA, z. B. unmittelbar nach einer Trennung (Sbarra & Ferrer, 2006) – <i>Nachweis von emotionalen Dysregulationen</i> bei Borderline-Patienten in der natürlichen Umgebung mittels PDA-Gerät (Reisch, Ebner-Priemer, Tschacher, Bohus & Linehan, 2008)
Atmung	– <i>Pneumatographie, Oximetrie, Capnographie</i> , z. B. Bestimmung des nasalen pCO <sub>2</sub> -Gehaltes als Korrelat zur Hyperventilation (Alpers, Wilhelm & Roth, 2005) – <i>Akustische Messung</i> von Husten, Räuspern etc. (Matos, Birring, Pavord & Evans, 2007)
Leistung	<i>Felddiagnostik mittels PDA-Gerät:</i> – AMBU: Ambulatory Monitoring and Behavior-Test-Unit (Buse & Pawlik, 2001) – Battery of Standardized Tests for Research with environmental Stressors (STRES; Psychology Software Tools, 1998) – Performance Evaluation Tests for Environmental Research (PETER; Kennedy & Bittner, 1980) – Einsatz eines ambulanten PDA-Stroop-Tests zur Analyse des Einflusses der „state anxiety“ (Waters & Li, 2008)
Substanzkonsum	– <i>Messung des Tabakkonsums</i> mittels PDA oder Interactive Voice Response (IVR) (McDaniel, Benson, Roesener & Martindale, 2005; Whalen, Jamner, Henker & Delfino, 2001; Shiffman, Kirchner, Ferguson & Scharf, 2009) und der Determinanten der Nikotinabstinenz (Chandra, Shiffman & Scharf, 2007) – <i>Messung des Alkoholkonsums (Trinkverhalten und Selbstkontrolle)</i> mittels PDA oder IVR: Messung mittels IVR (Searles et al., 2000); Messung mittels Handy-IVR (Collins et al., 2003); Messung mittels PDA (Muraven, Collins & Shiffman, 2005) – <i>Messung der pharmakologischen Compliance</i> mittels automatisiertem Medikamentenspender oder PDA (Dunbar-Jacob, Sereika, Rohay & Burke, 1998; Burke, 2001) – <i>Erhebung von „polydrug use“</i> in Abhängigkeiten von Tageszeiten mittels PDA (Hopper, Su & Looby, 2006)

Tabelle 1 (Fortsetzung):

Psychische Konstrukte, für die Strategien des ambulanten Assessments verwendet werden können

Konstrukt	Methode/Beispielstudie
Essen und Essstörungen	<ul style="list-style-type: none"> <li>– <i>Ambulante Erhebung physiologischer Korrelate</i> (Delichatsios et al., 2001)</li> <li>– <i>Messung des Essverhaltens mittels DietMate</i> (www.dietmate.com) und relevanter Prädiktoren wie Stress und Affekt (E. T. Barker, Williams &amp; Galambos, 2006; Burke, 2001)</li> <li>– <i>Erfassung des nächtlichen Essverhaltens</i> beim Nacht-Ess-Syndrom mittels EMA Protokoll (Boseck et al., 2007)</li> </ul>
Stress, strain und workload	<ul style="list-style-type: none"> <li>– <i>Messung physiologischer Korrelate</i>: Ambulante Messung in der Arbeitswelt (Hanson &amp; Barzanski, 2001; Myrtek, 2004; Rau, 2001)</li> </ul>
Psycho-pathologie	<ul style="list-style-type: none"> <li>– <i>Messung der Symptome bei Zwangsstörungen</i> mittels IVR als Selbstassessment (Bachofen et al., 1999)</li> <li>– <i>Messung von Kaufimpulsen und Stimmung bei Kaufsucht</i> mittels PDA (Silbermann, Henkel, Müller &amp; de Zwaan, 2008)</li> <li>– <i>Messung psychotischer Symptome und Stimmung bei Schizophrenen</i> mittels PDA (Kimhy, Delepaul &amp; Corcoran, 2006)</li> <li>– <i>Messung des Spielverhaltens und der vorübergehenden, begleitenden und nachfolgenden Stimmung</i> mittels IVR (Gee, Coventry &amp; Birkenhead, 2005)</li> </ul>
Gratifikationsmessung	<ul style="list-style-type: none"> <li>– <i>Erhebung physiologischer Parameter</i> mittels mobiler Geräte in der Medienpsychologie (Schlütz &amp; Scherer, 2001)</li> </ul>
Gesundheitspräventives Verhalten	<ul style="list-style-type: none"> <li>– <i>Erhebung der Sonnenexpositionsdauer</i> mittels ambulanten UV-Messgeräten und Analyse von Einflussfaktoren mittels elektronischen Tagebuch (Thieden, 2008)</li> </ul>
Soziale Interaktion	<ul style="list-style-type: none"> <li>– <i>Erfassen von paraverbalen</i> (z. B. Lachen) und <i>verbalen Interaktion</i> mittels „Electronically Activated Recorder (EAR)“ (Mehl &amp; Holleran, 2007)</li> </ul>

### 3 Anwendungsbegründung

Die Vorteile der Methoden hängen mit der Möglichkeit der Messung im Felde zusammen. Damit wird die *Reaktivität der Labormessung* eingeschränkt. Folgende Metapher von Suls und Martin bringt diesen Aspekt treffend auf den Punkt: „Although a pressed flower in an herbarium is useful to the botanist, it is not the same as a flower in bloom. Observing the flower in the wild, ‚in situ‘, provides supplementary, if not essential, information“ (Suls & Martin, 1993, S. 3). Wie bedeutsam die Messung *in situ* insbesondere im Bereich der Klinischen Psychologie und der Medizin ist, macht das Phänomen des „office hypertension“ (auch „white coat hypertension“ genannt) deutlich. In vielen Forschungsarbeiten (z. B. Pickering, Harshfield, Deveraux & Laragh, 1985) konnte gezeigt werden, dass die Messung des Blutdrucks in der Arztpraxis oder im Krankenhaus meist zu höheren Werten führt als bei der Blutdruckmessung in der häuslichen Umgebung. Dies kann falsche therapeutische Konsequenzen

nach sich ziehen. Dieses Phänomen konnte erst mit ambulanten Messmethoden aufgedeckt werden. Inzwischen gibt es eine Vielzahl von Studien, die zeigen, dass der prädiktive Wert einer zeitnahen ambulanten Blutdruckmessung im Hinblick auf kardiovaskuläre Folgeerkrankungen höher ist als die einmalige Messung in der Arztpraxis (z. B. Eguchi et al., 2008).

Eine Messung in der natürlichen Umwelt der Personen ermöglicht es, die Einflüsse von Situationen, persönlich gewählten Kontexten und deren Interaktion zu erfassen, also „capturing life as it is lived“ (Bolger, Davis & Rafaeli, 2003). Insbesondere dann, wenn Personen nicht wiederkehrend die psychologischen Einrichtungen besuchen können, bietet sich die ambulante Messung an. Die Messung kann dann automatisiert in der gewohnten Umgebung der Personen und ohne Testleiter im Feld erfolgen. Als Beispiel ist die Erfassung des Drogenkonsums bei Obdachlosen zu nennen, die mittels Mobiltelefon standardisiert zum aktuellen Substanzmissbrauch, zur Stimmung und zu Symptomen befragt werden können (z. B. Freedman, Lester, McNamara, Milby & Schumacher, 2006). Als weiteres Beispiel sei eine Studie genannt, die die Dynamik des emotionalen Erlebens unmittelbar nach dem Geschlechtsverkehr erfasst. Es ist offensichtlich, dass dies nur in der natürlichen Umgebung der Testpersonen ohne Anwesenheit eines Testleiters möglich ist und wiederholte Messungen mit entsprechenden PDA-Geräten sinnvoll sind (Shrier, Shih, Hacker & de Moor, 2007).

Durch die gleichzeitige Erfassung des Kontextes kann dessen Einfluss auf relevante psychologische Variablen erfasst werden. Zum Kontext gehört auch die Beeinflussung des sozialen Umfelds. Bei gleichzeitiger Datenerfassung mehrerer Personen mittels ambulanter Messmethoden kann zusätzlich die Dynamik der Interaktion analysiert werden, beispielsweise die gegenseitige Beeinflussung der Stimmung je nach persönlicher Zielerreichung in einer Beziehung (Meegan & Goedereis, 2006) oder in Abhängigkeit von partnerschaftlichen Konfliktlösungen (Merrilees, Goeke-Morey & Cummings, 2008). Dabei ist auch die Erfassung der sozialen Interaktionen durch direkte Aufzeichnung der verbalen und nonverbalen Äußerung möglich (z. B. Mehl, Gosling & Pennebaker, 2006).

Einen weiteren Vorteil stellt die *direkte und unverzerrte Messung* dar. Während im Fragebogen Angaben zum Verhalten meist selbst- oder fremdreflektiert und retrospektiv konstruiert sind (Brewin, Andrews & Gotlib, 1993), erlaubt das ambulante Assessment die unmittelbare Messung von Verhaltensmaßen, die Messung von „hot cognition“ und „hot emotion“ (Schwartz & Sudman, 1994). Durch die Notwendigkeit, beim Ausfüllen des Fragebogens gedanklich einen Blick in die Vergangenheit zu werfen, kommt es zu typischen Verzerrungen. Dieser sogenannte „retrospection bias“ ist mehrfach repliziert worden (vgl. DeLongis, Hemphill & Lehmann, 1992; Hedges, Krantz, Contrada & Rozanski, 1990; Käßler, Brünger & Fahrenberg, 2001; Pohl, 2004).

Eine Vielzahl von *Verzerrungstendenzen* sind bei der Messung durch Fragebögen beschrieben worden: Bezüglich der Stimmung- und Befindlichkeitsmessung zeigt sich bei Fragebögen, die auf vergangene Zeiträume als retrospektive Selbstbeobachtungsfenster Bezug nehmen, häufig ein negativer Retrospektions-effekt, das heißt, das Befinden und die Symptome werden negativer geschildert als sie aktuell erlebt werden. Die Aufdeckung eines solchen Effektes stellt selbst eine wichtige diagnostische Größe dar. So zeigt sich, dass die Stärke des negativen Retrospektions-effektes mit der emotionalen Labilität – operationalisiert über das Freiburger Persönlichkeitsinventar (FPI-R) – der Testperson im Zusammenhang steht (Käppler & Rieder, 2001). Ebner-Priemer et al. (2006) konnten zeigen, dass Personen mit Borderline-Störung im Hinblick auf die Messung der Stimmung einen negativen Recall-Bias haben, während nicht erkrankte Personen einen positiven Recall-Bias haben. Im Vergleich zu einer unmittelbaren ambulanten Messung zeigen depressive Personen bei der Einschätzung der Stimmung eine generelle Intensivierung des Affekts bei einer retrospektiven Einschätzung (Ben-Zeev, Young & Madsen, 2009)

Weiterhin konnten Stone et al. (2000) zeigen, dass die retrospektive Schmerzerfassung durch die Salienz des vorherigen stärksten Schmerzes und durch die Zeitdauer zwischen Auftreten und Erfassung (Recency-Effekt) mitbestimmt wird. Kikuchi und Kollegen (2006) konnten in einer Studie an Personen mit Spannungskopfschmerz zeigen, dass es insbesondere bei starker Variabilität der Schmerzen einen großen Unterschied zwischen unmittelbarer und verzögerter Messung gibt. Offensichtlich sind Personen nur schlecht in der Lage, eine Vielzahl von Stimmungen und Schmerzzuständen über einen längeren Zeitraum zu erinnern. Daher ist auch eine verlässliche Mittelwertbildung auf statistischem Niveau über zurückliegende Zeiträume („Wie stark waren die Schmerzen in den letzten 24 Stunden?“) kaum möglich. Sollen zusätzlich zur Symptomwahrnehmung auch noch die Variabilität und beeinflussende Faktoren (Stress, Stimmung) erhoben werden, erhöhen sich die Differenzen zwischen unmittelbarer und verzögerter Einschätzung (Gloster et al., 2008).

Fehler in den Angaben der Testpersonen ergeben sich auch durch die falsche Interpretation von Zeiträumen in klassischen Fragebögen. Bailey, Martin, Lynch und Pollock (2000) zeigten, dass die Instruktion an die Testpersonen, die Schmerzen im letzten Monat zu bewerten, nicht im Sinne der zurückliegenden 30 Tage interpretiert wird, sondern meist im Sinne des letzten Kalendermonats. Die in Fragebögen häufig anzutreffende Mittelwertbildung („Wie müde waren Sie im Durchschnitt während der letzten zwei Wochen?“) ist ebenfalls systematischen Verzerrungen unterworfen. Der Mittelwert mehrerer EMA-Messungen korreliert nur schwach mit der retrospektiven Mittelwertbildung der Testpersonen (vgl. Stone et al., 2000). Verzerrend ist auch, dass der Abruf von Erlebnissen und Erfahrungen zum Zeitpunkt *t1* durch spätere Er-

fahrungen zum Zeitpunkt *t2* beeinflusst wird („effort-after-meaning“, vgl. Stone et al., 1999).

Die Passung zwischen der Stimmung bei der Abfrage und der relevanten Stimmung, auf die sich der Fragebogen bezieht, beeinflusst die Intensität der Angaben (Salovey, Sieber, Jobe & Willis, 1993). Weiterhin gibt es Verzerrungen bei der Abfrage von Zeiträumen („Wie lange haben Sie nicht mehr geraucht?“) und von Zeitpunkten („Wann hatten Sie die letzte Panikattacke?“). Beispielsweise wird der Abstand zu kritischen Events verkürzt wahrgenommen („telescoping“; vgl. Schwartz & Sudman, 1994). Zur Einschätzung von Häufigkeiten werden bei geringen Häufigkeiten Ereignisse gezählt, bei hoher Häufigkeit werden dagegen eher Schätzungen vorgenommen, die dann wiederum durch typische Verzerrungen beeinflusst sind.

Die angesprochenen Verzerrungen tauchen am häufigsten bei globalen Selbstschilderungen auf („Wie häufig trinken Sie durchschnittlich Alkohol?“), in geringem Ausmaß beim episodischen Abruf („Wie viel Alkohol haben sie bei der letzten Feier getrunken?“). Den geringsten „recall bias“ findet man bei unmittelbaren Erhebungen, wie es beim EMA oder ambulanten Assessment der Fall ist (Shiffman, 2000).

Das ambulante Messen verfolgt also das Ziel, die Tücken klassischer Fragebögen zu überwinden. Während zur Messung von Einstellungen und Einschätzungen kein Weg an der klassischen Fragebogenmethode vorbei führt, kritisieren Fahrenberg et al. (2007) den Fragebogen als fehlerhafte „Ersatzmethode“, „wenn es in Wahrheit nicht um mentale Repräsentation, sondern direkte Erfassung von Erleben und Verhalten geht“ (S. 14). Pawlik (1996) hinterfragt in seinen Ausführungen, warum die Verhaltensbiologie sich der Beobachtung im Felde bedient, während in der Psychologie als Verhaltenswissenschaft interindividuelle Unterschiede inmitten von intraindividuellen Konsistenzen (Laborumgebung) erhoben werden. Er fordert daher: „one should better (and more economically) try to carry the assessment methodology into the ‚field‘ – and not attempt to carry the ‚field‘ into the laboratory“ (Pawlik, 1996, S. 23). Dies sichert die *ökologische Validität* der Daten: „In this sense, ambulatory assessment carries differential psychology into the ‚real world‘“ (Pawlik, 1996, S. 23).

Natürlich sind auch bei der unmittelbaren Messung im Felde Verzerrungen möglich. So lässt sich die Tendenz zur positiven Selbstdarstellung nicht ausschließen. Durch die Unmittelbarkeit der Erhebung werden jedoch *die* Verzerrungen ausgeschlossen, die mit der Notwendigkeit zur Rekonstruktion von Erinnerungen zu tun haben: „This emphasis on episodic versus semantic knowledge usually allows momentary self-reports greater accuracy for the details surrounding particular events or brief time periods, as they are not as tinged with the

identity of the person completing the self-report“ (Beal & Weiss, 2003, S. 444). Entsprechende Rekonstruktionen mögen bei Persönlichkeitsfragebögen intendiert sein, im Hinblick auf Verhaltens- und Erlebensaspekte (Stimmung, Schmerz etc.) können diese jedoch zu Unter- oder Überschätzung führen, aus denen sich dann falsche therapeutische Konsequenzen ergeben.

Ein weiterer Vorteil dieser Messstrategie ergibt sich durch die Möglichkeit der *kontinuierlichen Erfassung* von Daten. Immer dann, wenn die Within-subject-Varianz und prozesshafte Veränderungen von Interesse sind, haben solche Verfahren ihre Berechtigung. Der Fokus auf intrapersonelle Schwankungen ermöglicht es, die Wertigkeit interpersoneller Differenzen zu bewerten. So zeigt eine Untersuchung von Fleeson (2001), der die „Big Five“ über 2 bis 3 Wochen mittels EMA erhob, dass die Within-subject-Variabilität größer ist als die Variabilität zwischen den Personen. Für die Konstrukte „Conscientiousness“ und „Extraversion“ war sie sogar höher als die Between-subject-Variabilität. Die intraindividuelle Variabilität ist durch jeweilige Situationen mitbestimmt und damit mehr als ein unsystematischer Fehler.

Dass die interindividuellen Unterschiede in den intraindividuellen Schwankungen selbst einen diagnostischen Wert haben, zeigen Riis et al. (2005). Sie untersuchten den Stimmungsverlauf von Gesunden im Vergleich zu Dialysepatienten und konnten mittels EMA verdeutlichen, dass sich beide Gruppen in ihrem Stimmungsverlauf unterschieden: Patienten zeigen im Vergleich zu Gesunden eine „hedonistische Adaptation“, d. h. eine zunehmend bessere Einschätzung der Stimmung. Die einmalig und retrospektiv eingeschätzte Lebenszufriedenheit korrelierte mit den Mittelwerten der EMA-Messung stärker bei den Patienten als bei den Gesunden. Insbesondere dann, wenn es um die Verlaufsmessung geht und längere Zeiträume erfasst werden sollen, die mit der klassischen Laborforschung kaum möglich sind, ist das ambulante Assessment die Methode der Wahl. Mattocks und Kollegen (2007) beschreiben beispielsweise eine automatisierte Messung der Beweglichkeit über ein ganzes Jahr, mit der auch saisonale Schwankungen erfasst werden konnten.

Durch die unmittelbare Erfassung („momentary assessment“) ist es mit den technischen Geräten auch möglich, die Reaktion auf seltene Ereignisse zu erfassen, die sonst nur retrospektiv und mit entsprechender Verzögerung erfasst werden können. So können beispielsweise das Erleben von Rückfällen beim Alkoholmissbrauch (McKay, Franklin, Patapis & Lynch, 2006) oder das Erleben in bestimmten sensiblen Etappen der Nikotinentwöhnung erfasst werden (Shiffman, Scharf & Shadel, 2006).

Die EMA-Methodik erlaubt auch die *gleichzeitige Erfassung verschiedener Variablen*. So können z. B. physiologische Daten mit den parallel erhobenen Bewe-

gungsparametern oder der selbstbewerteten Stimmung verglichen werden. Hiermit können Reiz-Reaktions-Kontingenzen aufgedeckt werden. Goldberg und Kollegen (2007) gelang es beispielsweise, durch die gleichzeitige Erfassung von prämenstruellen Beschwerden (PMS) und Kopfschmerzen, entsprechende Koinzidenzen und bestimmte Schmerzmuster zu identifizieren. Ambulantes Assessment kann somit Within-Prozesse entschlüsseln helfen, die sich in interindividuellen Differenzen nicht zeigen.

Durch die automatisierte Auswertung der Daten ist auch eine *direkte Rückmeldung* an die Testpersonen möglich. Daher spielt das ambulante Assessment insbesondere im Bereich des „self-monitorings“ eine wichtige Rolle. *Self-monitoring* ist definiert als „any behavior that entails reflecting on and providing self-reports about a specific behavior, such as drinking, or emotional state regardless of the mode of reporting (e. g., via paper-and-pencil diaries, electronic pagers or palm pilots, telephone data capture systems)“ (Simpson, Kivlahan, Bush & McFall, 2005, S. 241). Nicht jedes „self-monitoring“ ist auch ein „ambulatory monitoring“, sondern definiert sich zunächst nur durch die erhöhte Aufmerksamkeit auf das Selbsterleben oder Verhalten. Erst wenn diese Selbstwahrnehmung im Feld (ambulant) und unmittelbar („momentary“) stattfindet, sind die Kriterien des ambulanten Assessments erfüllt.

Die Möglichkeit der direkten Rückmeldung der selbst eingegeben Daten spielt besonders im Bereich der Patientenedukation eine wichtige Rolle. Als Beispiele sind hier die Trainings zur Symptomwahrnehmung bei Asthma- oder Diabeteypatienten (beim „hypoglycemia unawareness syndrom“) zu nennen (Schandry & Leopold, 1996) oder die automatische Kontaktierung von Schmerzpatienten durch den Psychologen in Abhängigkeit von der Selbsteinschätzung des Schmerzes (Jamison, Fanciullo & Baird, 2004). Mittels mehrfacher Zyklen von Messung und Intervention ist es möglich, adaptiv das Verhalten (z. B. das Suchtverhalten) von Personen an einen Sollzustand anzunähern. Leopold und Schandry (2001) sprechen dann von „computer-tailored interventions“. Werden nicht nur Diagnosen, sondern auch Handlungsempfehlungen aus der Messung abgeleitet, dann verschmelzen Diagnostik und Therapie zu sogenannten „*Theranostics*“. Die unmittelbare und unauffällige *Rückwirkung* der erhobenen Daten ohne Beteiligung auswertender Personen wird mit der Entwicklung von „Ambient-home“- oder „Smart-home“-Technologien (vgl. Abschnitt 8.3) ermöglicht.

#### 4 Technische Voraussetzungen

Ein ambulantes Assessment benötigt nicht zwingend technische Geräte. Die Mitgabe eines Fragebogens an Klienten mit der Bitte, diesen nach einem bestimmten Ereignis (z. B. nach einer Panikattacke) auszufüllen, stellt letztlich

auch schon eine Form des ambulanten Assessments dar. Technische Geräte werden erst dann benötigt, wenn ein elektronisch gesteuertes Signal zur Dateneingabe erfolgt (sog. *Prompting*) oder automatisiert Parameter erhoben werden. Mit der technischen Weiterentwicklung wird die Möglichkeit einer unmittelbaren Datenerfassung im Feld sicherlich ansteigen. Kasten 1 gibt einen Überblick über die derzeitigen technischen Mittel zur Umsetzung des ambulanten Assessments.

**Kasten 1:**

Beispiele für technische Geräte zur Realisierung des ambulanten Assessments

- ambulante physiologische Messvorrichtungen, z. B. „Vitaport“ und „LifeShirt“
- ambulante EEG-Messvorrichtungen (EEG: Elektroenzephalogramm)
- Handheld-Computer (PDA), inkl. Video- und Sprachaufzeichnungsfunktion
- RFID-Chips (RFID: Radio Frequency Identification; Kundenkarten, Zugangskontrollen, Tracking)
- Telefon und „cellular phones“ (Handys), inkl. WAP, GPRS, SMS, MMS, UMTS<sup>3</sup>
- automatisierter Anruf von Probanden mittels „Interactive Voice Response“ (IVR)
- Gruppenmessverfahren: „Voting machine“, Tele-Dialog (TED)
- „electronic surveillance measures“/„biomechanical activity devices“, inkl. Global Positioning System (GPS)

Für *physiologische ambulante Messungen* liegt inzwischen eine Vielzahl etablierter Verfahren vor. Eine Auflistung aktuell gebräuchlicher Mehrkanal-Rekorder für die ambulante physiologische Messung findet sich bei Ebner-Priemer und Kubiak (2007). Prominente deutsche Entwicklungen sind „Vitaport“ und „Varioport“. Inzwischen gibt es ergonomisches Equipment, bei dem das Anbringen von Elektroden entfällt. Das sogenannte „LifeShirt“ ist ein ärmelloses Hemd, das die Bewegung von Rippen und Bauch sowie weitere kardiopulmonare, physiologische Parameter misst. Das Hemd ist waschbar und kann komfortabel über einen längeren Zeitraum sowohl beim Schlafen als auch bei motorischen Aktivitäten getragen werden (Wilhelm, Roth & Sackner, 2003). In einer Vergleichsstudie zeigt sich eine ebenso exakte Messung wie bei stationären Geräten (Heilman & Porges, 2007). Ein weiteres wichtiges Messverfahren stellt die *automatisierte Erfassung der Bewegung* dar. Es gibt inzwischen eine große Vielfalt an Messgeräten, die entweder nur die Anzahl der getätigten Schritte (Pedometer) oder – umfangreicher – die körperlichen Bewegungen (Akzelerometer) in einer oder mehreren Bewegungsachsen messen. Diese werden vor allem in der Sport- und Gesundheitspsychologie eingesetzt, da sie gegenüber der Messung des Bewegungsverhaltens mittels Fragebögen eine genau-

<sup>3</sup> WAP: Wireless Application Protocol; GPRS: General Packet Radio Service; SMS: Short Message Service; MMS: Multimedia Messaging Service; UMTS: Universal Mobile Telecommunications System

ere Analyse ermöglichen (für eine Übersicht vgl. de Vries, Bakker, Hopman-Rock, Hirasig & van Mechelen, 2006). Beispielhaft kann hier eine Studie von McMurray et al. (2008) genannt werden, die verdeutlicht, dass übergewichtige Personen die Anzahl der täglichen Bewegungen eher überschätzen. Ein weiterer Einsatzort ist die Erfassung der Bewegungen bei Dementen, beispielsweise um den Erfolg medikamentöser und nicht medikamentöser Interventionen zu prüfen (Mahlberg & Walther, 2007; Reuschenbach, 2006). Eine automatisierte Erfassung mittels Akzelerometer ermöglicht die genauere Erfassung des agitierten oder herausfordernden Verhaltens als die Einschätzung durch Pflegende (Nagels, Engelborghs & Vloeberghs, 2007).

Neben dieser passiven Messung von physiologischen Daten hat sich ein umfangreicher Forschungsbereich zur *aktiven und reaktiven Messung* mittels technischer Geräte gebildet. Zur Präsentation von Fragebögen im Feld werden meist *mobile Handcomputer* (PDA oder Handheld-Computer) verwendet. Neuerdings sind auch Eingaben über die *Telefontastatur* des Handys möglich. Vorteile der Nutzung solcher Geräte sind: (a) die genaue Zeit der Dateneingabe kann erfasst werden, (b) zur Kontrolle der Eingabezeit können bestimmte Zeitfenster definiert werden, in denen eine Antwort erforderlich oder möglich ist, (c) die Latenzzeit zwischen „prompting“ und der tatsächlichen Eingabe kann erhoben werden und hat einen weitergehenden diagnostischen Wert, z. B. zur Messung der Compliance, (d) im Hinblick auf die Verlaufskontrolle ist es hilfreich, dass die vorherigen Eingaben für die Person nicht mehr verfügbar sind; bei den meisten PDA-Verfahren werden die bisherigen Dateneingaben abgespeichert und sind dann nur für die Testleiter zugänglich, (e) die Item-Reihenfolge kann auch über mehrere Erhebungen hinweg verändert werden, (f) es ist eine adaptive Testung möglich, wenn das Gerät genügend Kapazität für ein entsprechendes Vorgabeprogramm hat, (g) die Messung kann an die Bedürfnisse der Testpersonen angepasst werden, entweder durch Reduktion der Items oder auch durch die Möglichkeit, eine Messung auszusetzen, (h) es ist eine direkte Rückmeldung an die Testpersonen möglich. Für eine detaillierte Darstellung dieser Vorteile vgl. Bolger et al. (2003). Weiterhin kann die SMS-Funktion des Handys auch zur Sammlung qualitativer Daten im Feld genutzt werden (z. B. Roenkae et al., 2010).

Die heutigen PDA-Geräte und Smartphones sind meist mit folgenden *Funktionen* ausgestattet:

1. Ein *Display*, das zur Itemvorgabe und Instruktion genutzt wird
2. Verschiedene *Möglichkeiten zur Dateneingabe*, entweder über die Tastatur oder durch einen Touchscreen. Ob der Touchscreen mit dem Finger oder mit entsprechenden „Pens“ bedient wird, hängt von der Größe der Itemdarstellung und dem Gerät ab. Inzwischen sind auch Aufzeichnungen von verbalen Daten mittels PDA möglich. Hierbei können dann neben der In-

haltsanalyse auch linguistische Komponenten analysiert werden (Mehl, Pennebaker, Crow, Dabbs & Price, 2001).

3. Die Systeme sind heutzutage meist mit dem *Betriebssystem* PalmOS, Symbian, Android, Apple iOS oder Windows Mobile ausgestattet. Fragebögen werden meist am stationären Computer erstellt und dann auf die mobilen Geräte übertragen. Inzwischen gibt es zahlreiche kostenlose Software, mit der entsprechende Testinstrumente entwickelt werden können. Bekannte Beispiele sind: das *Purdue Momentary Assessment Tool* (PMAT) von Weiss, Beal, Barros und MacDermid (2003), das *Experience-Sampling Program* (ESP) von Barrett und Feldman-Barrett (2000), *MyExperience*, ein Software-Programm für Smartphones, das auch ein kontextspezifisches Sampling ermöglicht (<http://myexperience.sourceforge.net>) und *MONITOR*, ein Softwareprogramm, das vor allem in Deutschland eingesetzt wird (vgl. Käßler et al., 2001). Eine Übersicht über den Stand der kommerziellen und nicht kommerziellen Software- und Hardware-Entwicklung gibt der Beitrag von Ebner-Priemer und Kubiak (2007). Mithilfe der Programme lassen sich Items mit verschiedenen Antwortformaten erstellen, zum Beispiel offene Fragen, Multiple-Choice-Fragen, visuelle Analogskalen. Ebenso sind psychomotorische Aufgaben und adaptive Testformate möglich. Einen Vergleich der Software-Pakete mit unterschiedlichen Features liefern Le und Hat (2006).
4. Geräte haben die Möglichkeit zur *Datenspeicherung* oder zum direkten Transfer der Daten mittels drahtloser Technologien.
5. Die Geräte sind mit einem integrierten *Prompting-System* (auch „pager“ oder „beeper“ genannt) ausgestattet, das die Testperson durch ein Signal an die Eingabe von Daten erinnert. Möglich ist aber auch eine eventabhängige Eingabe, bei der die Person selbst den Zeitpunkt bestimmt.

Vor der Nutzung entsprechender Systeme muss sichergestellt sein, dass die Testperson den Umgang mit dem Gerät sicher beherrscht, denn im Feld ist nicht immer ein Ansprechpartner erreichbar, der die Bedienung erläutern kann. Die Eingabegeräte müssen leicht und – je nach Einsatz im Feld – auch robust sein (Flechter, Erickson, Toomey & Wagenaar, 2003; Palmblad & Tiplady, 2004). Neben dem Einsatz von gängigen Handheldgeräten gibt es auch Datenrekorder, die speziell für die Datenerfassung entwickelt wurden. Prominentestes Beispiel eines solchen „behavioral data monitor“ ist der AMBU-Rekorder (Ambulatory Monitoring and Behaviour-Test Unit; vgl. Buse & Pawlik, 1996, 2001).

Als *Testverfahren* oder *Messinstrumente* kommen meist Adaptationen klassischer Fragebögen oder selbst entwickelte Items mit Likert-Skalen zum Einsatz. Stone und Shiffmann (2002) empfehlen wegen der häufig kleinen Displays eine „Übersetzung“ klassischer Items in ein auf dem PDA lesbares Format. So kann beispielsweise bei der Präsentation einer siebenstufigen visuellen Skala das

Display zu klein sein für eine horizontale Anordnung (Palermo, Valenzuela & Stork, 2004).

Buse und Pawlik (1996) berichten von einer Vielzahl von „self-monitoring inventories“, die beispielsweise mit dem AMBU-Gerät erhoben werden können: Verhaltensitems („Was tun Sie gerade?“); Stimmungs- und Motivationslisten („Wie fühlen Sie sich gerade?“); Abfrage von körperlichen Beschwerden (Hunger, Schmerzen etc.). Weiterhin sind auch Testverfahren und Testbatterien zur Messung der kognitiven Leistungsfähigkeit integrierbar, z. B. der *Rota-Test* zur Messung der Wahrnehmungsgeschwindigkeit und des Arbeitstempos, der *Maze-Test* zur Messung des visuellen Kurzzeitgedächtnisses, eine Adaptation des *d2-Tests* zur Messung der Aufmerksamkeit oder ein *Tracking-Test* zur Messung der visuellen Aufmerksamkeit und psychomotorischen Koordination unter Zeitdruck (vgl. Buse & Pawlik, 2001).

Der besondere Nutzen solcher Datenrekorder liegt in der parallelen Erhebung verschiedener Variablen. Häufig finden sich die Abfrage der Symptome mittels PDA-Inventar und die gleichzeitige Erfassung der Physiologie. Eine Möglichkeit der parallelen Erfassung von Daten zur subjektiven Befindlichkeit mit physiologischen Parametern liefert das *Freiburger Monitoring System* (FMS; Myrtek, 2004; Stephan, Mutz, Feist & Weiß, 2001). Hiermit ist es beispielsweise möglich, die subjektive Symptomwahrnehmung eines Asthmatikers, die mittels „Patient Experience Diaries“ (PEDs) erhoben wurde, mit den gleichzeitig erfassten respiratorischen Werten zu vergleichen. Der Faktor Symptomwahrnehmung stellt eine eigene diagnostische Variable dar. Ein weiteres Beispiel ist die Messung von Bewegungsdaten mittels Akzelerometrie und der parallelen Erhebung von Migräneattacken. So lassen sich Ursache-Wirkungs-Fragen und Zeitverläufe klären (Tulen, Stronks, Bussmann, Peplinkhuizen & Passchier, 2000).

Eine andere Form des ambulanten Assessments ist die Nutzung von „Telecommunication Monitoring Systems“ (TMS) oder auch „Interactive Voice Response“ (IVR) genannt. Anstelle eines schriftlich präsentierten Inventars erfolgt die Dateneingabe über ein Telefon. Die Anwahl des Telefoncomputers kann automatisiert erfolgen oder in der Verantwortung der Testperson liegen. Die Eingabeaufforderung und die Datenanfrage sind durch eine aufgezeichnete Stimme oder durch einen Interviewer möglich. Die reale Anwesenheit eines Telefoninterviewers schränkt jedoch möglicherweise die Compliance gegenüber einer elektronischen Abfrage insbesondere bei der Aufdeckung von Problemverhalten ein (Turner et al., 2005). Eingaben sind entweder über die Tastatur des Telefons möglich (Ritterband et al., 2001) oder es erfolgt die Aufzeichnung verbaler Daten. Freie Angaben ermöglichen neben der inhaltlichen Auswertung zusätzlich auch eine Analyse paraverbalen Daten, die z. B. auf die Stimmung der

Person hinweisen. Es können Festnetzanschlüsse in der häuslichen Umgebung oder auch Mobiltelefone genutzt werden. Das Prompting kann auf zwei Wegen erfolgen: Entweder wird die Testperson automatisiert angerufen und die Befragung beginnt sofort. Oder ein externes Geräte („beeper“) bzw. das Telefon selbst senden einen Hinweis, der dazu auffordert, sich in das Voicemail-System einzuwählen (Collins, Kashdan & Gollnisch, 2003). Die Anwahl des IVR kann aber auch von der Person selbst gesteuert werden, wenn es sich um ein „Event-Sampling“ handelt (Gee, Coventry & Birkenhead, 2005). Vor der eigentlichen Dateneingabe wird der Proband gebeten, sich durch einen Sicherheitscode zu legitimieren.

IVR-Systeme wurden bisher zur Messung des Trinkverhaltens (Searles, Helzer & Walter, 2000), der Spielsucht und der damit in Zusammenhang stehenden Stimmung (Gee et al., 2005), zur Messung von Zwangsstörungen (Bachofen et al., 1999) und zur Analyse des Essverhaltens (Delichatsios et al., 2001) eingesetzt. Für eine Überblicksarbeit mit dem Schwerpunkt auf dem Bereich der medizinischen Diagnostik vgl. Corkey und Parkinson (2002).

Vorteile gegenüber der PDA-Methode sind die geringeren Kosten, die unauffällige Integration in den Alltag und der geringe Aufwand für das Training zur Dateneingabe. Für den Einsatz ist lediglich ein Computer mit Internetanschluss und entsprechende Software notwendig (z. B. „ProPhone“ von [www.bingsoftware.com](http://www.bingsoftware.com)). Der Einsatz bei klinischen Stichproben zeigt, dass etwa 75 % der Aufforderungen zur Dateneingabe beachtet werden. Verbesserungen sind durch gezielte Stichprobenbildungen, durch Anreize zur Teilnahme von der Testperson und durch wiederholte Erinnerungen zu erreichen. Hinsichtlich der Messreaktivität gibt es widersprüchliche Befunde, die durch die Stichprobe, die Untersuchungsabsicht und die Häufigkeit der Messung begründet sind (vgl. Simpson et al., 2005).

Mit der gegenwärtigen Entwicklung von „Smartphones“, bei denen die Handfunktion mit den Möglichkeiten eines mobilen Computers verschmilzt, sind zukünftig sicherlich erweiterte Erhebungsmethoden möglich. Denkbar ist, dass das Prompting per Anruf erfolgt und eine Messung mittels komplexer Items (z. B. Videos oder psychomotorische Aufgabe) abläuft.

Bezogen auf das Reizmaterial sind nun auch Bilder und Videos als Items möglich (Intille, Rondoni, Kukla, Iaconi & Bao, 2003). Neben der klassischen Auswahloption von Antworten können auch Sprachaufzeichnungen (Mehl et al., 2001) sowie Videoaufnahmen durch die Testperson selbst erfolgen (Intille et al., 2003). Die Bewertung eines solchen Videos kann von der Testperson selbst vorgenommen werden. Dabei dient die Video- oder Sprachaufzeichnung, die das Gerät selbstständig vornimmt, selbst als Bewertungsgrundlage („Wie

haben Sie sich in dieser Situation gefühlt?“). Dies ermöglicht eine Datenerhebung in zwei Stufen: Einmal als Aufzeichnung des Verhaltens und dann als Präsentation dieser Aufzeichnung quasi als „Item“ für eine nachfolgende Bewertung durch die Testperson. Für die Verhaltensanalyse hat eine solche gestufte Erhebung den Vorteil, dass der Verhaltensfluss nicht unterbrochen wird.

Neben der zeit-, signal- oder eventgesteuerten Dateneingabe gibt es erste Versuche einer kontextsensitiven Abfrage. Das Prompting erfolgt beispielsweise nur dann, wenn man ein Einkaufszentrum betritt oder bestimmten Personen begegnet. Eine entsprechende Aktivierung des Gerätes erfolgt durch ein GPS-Plug-In oder durch die Detektion eines RFID-Chips (Intille et al., 2003). Dabei sind ethische und rechtliche Richtlinien zu beachten (vgl. Abschnitt 7.4)

## 5 Methodische Aspekte

Die ambulante Messung und das ambulante Monitoring stellen besondere Anforderung an die Stichprobenbildung, das sogenannte „Sampling“ (vgl. Abschnitt 5.1), und an die Analyse von Varianzkomponenten in den Daten (vgl. Abschnitt 5.2). Durch die besonderen methodischen Anforderungen an die Analyse von Testdaten hat die Methode des ambulanten Assessments auch wesentlich zur Weiterentwicklung neuer statistischer Methoden beigetragen, die in anderen Kontexten, z. B. bei clusterrandomisierten Studien, zum Einsatz kommen können.

### 5.1 Sampling

Während die klassische Testdiagnostik davon ausgeht, dass bestimmte Traits ein Verhalten bedingen und damit das durch Items provozierte Verhalten ein Beispiel für das dahinter stehende Konstrukt ist, steht beim ambulatory assessment das Verhalten auch in Abhängigkeit vom Umfeld im Mittelpunkt. Generalisierungen sind nur auf dem Hintergrund gleicher Kontexte über eine Person hinweg möglich (vgl. Stemmler, 1996). Für die Datenerhebung sind daher die Erhebungssituationen vorher zu spezifizieren. Der Zeitpunkt, die Zeitdauer und der Auslöser der Messung müssen ausgehend von theoretischen oder praktischen Erwägungen geplant werden, um die Repräsentativität der Daten zu sichern. Sampling bezieht sich also nicht nur auf die Auswahl der Testpersonen, sondern auch auf die Zeiten und Situationen, in der die Datenerfassung stattfinden soll.

Geht es beispielsweise darum, das Bewegungsverhalten im Alltag zu messen, so sind entsprechende Aktivitätsmuster zwischen Wochenende und Werktagen

bei der Stichprobenbildung zu berücksichtigen. Soll das Schmerzniveau im Laufe eines Tages bestimmt werden, dann sind für die Stichprobenbildung die Veränderungen des Schmerzniveaus über den Tag hinweg und eventuelle schmerzleichternde Einflussfaktoren zu beachten. Erst anhand von repräsentativen Erhebungssituationen kann letztlich ein geeigneter Mittelwert gebildet werden, der einen Eindruck vom durchschnittlichen Tagesschmerz ermöglicht. In der EMA-Literatur werden bisher drei verschiedene Erhebungskontexte (Samplings oder recording paradigmas) unterschieden (vgl. Hufford et al., 2001; Suls & Martin, 1993):

1. *Event-contingent recording*: Hierbei bestimmt die Anwesenheit eines Events (z. B. eine Panikattacke) die Datenerhebung. Eine einfache Form ist die Dokumentation von Zeit, Ort und Setting eines relevanten Ereignisses, zum Beispiel eines Migräneanfalls. Nicht zwingend müssen mehrere Messzeitpunkte vorliegen. Auch schon ein seltenes Ereignis, das erfasst werden soll, reicht aus, um vom EMA zu sprechen. Ein „Ereignis“ wird entweder von der Testperson selbst wahrgenommen, oder es findet eine automatische Aktivierung ausgehend von physiologischen Messungen statt. In einer frühen Form des event-contingent recording von Aserinsky und Kleitman (1953) wurden beispielsweise kontinuierlich beim Schlaf Augenbewegungen aufgezeichnet. Sobald eine kritische Größe überschritten wurde, wurden die Personen geweckt und gebeten, Angaben zu ihren Träumen zu machen. In der Stressforschung wird das Überschreiten einer kritischen Herzfrequenz zum Anlass genommen, Angaben zum aktuellen Stresserleben zu erheben. Umgekehrt kann auch ein bewusst wahrgenommenes Event (z. B. eine Panikattacke) von der Testperson zum Anlass genommen werden, die psychophysiologische Messung zu starten. Das Wechselspiel verschiedener Messmethoden wird „interaktives Monitoring“ genannt. Ein solches Event-Sampling bringt folgende Probleme mit sich: (a) Die Möglichkeit zur Beantwortung von Fragebögen ist bei bestimmten Events nicht immer zeitnah möglich, z. B. bei Panikattacken oder Migräneanfällen. (b) Bei seltenen Ereignissen sind lange Aufzeichnungsperioden notwendig. (c) Liegt die Aufgabe für die Event-Beobachtung bei der Person, dann sind vorab genaue Absprachen darüber notwendig, welche Events als relevant erachtet werden und wie diese definiert sind. Das Verständnis von Events zwischen Testleiter und Testperson muss abgeglichen werden.
2. *Interval-contingent recording* (manchmal ist auch die Rede von „time-based-designs“): Hierbei findet in regelmäßigen („fixed-time schedule“) oder randomisierten Zeitabständen („random-time schedule“) eine Messung statt. Vorab müssen die genauen Zeiträume der gesamten Erhebung, die Zeitpunkte der Messungen, das Zeitfenster, auf das sich die Messung bezieht, und das Zeitfenster, in dem Eingaben möglich sind, genau spezifiziert werden. Eine empirische Erfassung des optimalen Zeitintervalls für besondere

Personenstichproben ist anzuraten (z. B. Ebner-Priemer & Kubiak, 2007). Zur Erfassung zyklischer Schwankungen müssen die Zeitfenster so gewählt sein, dass nicht nur besondere Ausprägungen (z. B. Leistungsspitzen) erfasst werden (vgl. Shiffman & Stone, 1998; Stone & Shiffman, 2002). Grundsätzlich gilt, dass mit zunehmender zeitlicher Entfernung von der momentanen Messung die oben angesprochenen Erinnerungsverzerrungen auftreten können. Daher sind zeitnahe Aufzeichnungen zu bevorzugen. Wird von der Person eine regelmäßige Eingabe verlangt, dann kann diese Aufforderung auch im Konflikt mit gleichzeitig ablaufendem Verhalten (z. B. Einkauf, Konzertbesuch etc.) stehen. Ein weiteres Problem, das bei fixierten Zeitabständen auftaucht, ist die Antizipation der Messung, was zu einer erhöhten Aufmerksamkeit im Hinblick auf die relevanten Variablen führen kann. Hierdurch wird die Symptomwahrnehmung verstärkt; dies führt zur Reaktivität der Messmethode (vgl. Abschnitt 7.2).

3. *Signal-contingent recording*: Bei dieser Stichprobenbildung erfolgt die Messung erst auf ein Signal hin. Ein solches Prompting geschieht meist durch das gleiche Gerät, auf dem die Fragebögen auszufüllen sind, durch das Telefon oder durch einen Pager. Das signal-contingent recording ist die Methode, die beim ESM verwendet wurde (vgl. DeVries, 1992; Larson & Csikszentmihalyi, 1983). Die Verteilung der Signale und damit der Messungen kann theoriegeleitet (z. B. zu einer bestimmten Tageszeit aufgrund zyklischer Schwankungen) oder zufällig erfolgen, um eine repräsentative Stichprobe zu erzielen. Zur Bildung einer repräsentativen Sammlung sind geschichtete Stichproben, Zufallsstichproben oder massierte Erhebungen zu bestimmten Events hin möglich.

Ergänzend zu diesen drei Sampling-Methoden ist das *kontinuierliche Recording* zu nennen, das meist beim ambulanten physiologischen Messen verwendet wird. Hierbei werden innerhalb eines definierten Zeitabschnittes ständig Daten übertragen. Beim *kontrollierten strukturierten Recording* (de Vries et al., 2006) erfolgt die Messung der physiologischen Daten in bestimmten vorher definierten Umwelten, z. B. am Arbeitsplatz. Dies ermöglicht auch die Aufdeckung von interindividuellen Unterschieden ausgehend von ähnlichen Situationen. Eine neuere Technik ist das *kontextsensitive Sampling*. Das Prompting wird dabei durch Kontextfaktoren (Räume, Personen) initiiert. Sensoren messen die Anwesenheit entsprechender Schlüsselsensoren und aktivieren dann ein entsprechendes Prompting (Intille et al., 2003).

Es sind Mischdesigns zwischen allen fünf Sampling-Methoden möglich. Als Beispiel sei hier das Sampling beim Einsatz des *Freiburger Monitoring Systems* (FMS) erläutert (vgl. Myrtek, 2004). Die teilnehmenden Probanden wurden durch ein Signal aufgefordert, Angaben zur Befindlichkeit und zu Aktivitäten zu machen, wenn physiologische Parameter („additional heart rate“) einen be-

stimmten kritischen Wert überschritten haben (event-contingent recording). Zusätzlich gab es zum Vergleich aber auch Eingabeaufforderungen außerhalb physiologischer Belastungsphasen, die zufällig oder zeitgesteuert initiiert waren („random feedback“). Hiermit sind Vergleiche der Abhängigkeiten der einzelnen Variablen sowohl mit als auch ohne Events möglich.

Die Wahl der Sampling-Strategie wird durch die Fragestellung, das Erhebungsdesign, die Belastbarkeit der Person, durch die relevanten psychologischen Konstrukte, durch den Erhebungszeitraum, die Art des Datenrekorders und die Prompting-Möglichkeiten bestimmt (Shiffman & Stone, 1998). Meistens werden beim signal-contingent recording 5 bis 25 Messungen pro Tag verwendet. Unabhängig von der Art des Samplings müssen dabei folgende Aspekte beachtet werden:

1. Da eine mehrmalige Messung notwendig ist, muss die Dateneingabe möglichst leicht von der Hand gehen, d. h. komplexe Antwortformate und lange Eingabezeiten sind zu vermeiden. Adaptive Testgestaltungen und ein vorheriges Training können hier hilfreich sein.
2. Durch die mehrmalige Darbietung identischer Items zu unterschiedlichen Zeitpunkten kann sich ein spezieller Antwortstil bilden. Auch hier sind adaptive Testgestaltungen oder randomisierte Itemabfragen hilfreich.
3. Zur Vermeidung von Antworttendenzen sollten die vorherigen Eingaben für die Person nicht einsehbar sein. Dies ist ein besonderer Vorteil der PDA-Methode im Vergleich zu Papier-Bleistift-Abfragen.
4. Auch wenn es sich um „momentary assessments“ handelt, müssen die Zeitfenster beachtet werden. Der Verweis auf „wie fühlen Sie sich gerade *jetzt* oder *momentan*“ ist wenig brauchbar, da sich die Angaben damit auf die Stimmung während der Dateneingabe bezieht. Es sollten kurze Zeitfenster (z. B. 5 Minuten) gewählt werden.

## 5.2 Auswertung

Die große Anzahl an längsschnittlichen Daten über Personen und Situationen hinweg stellt an die Auswertung große Anforderungen. Zur Analyse der gewonnenen Daten gibt es – ausgehend von der Fragestellungen und der Art der Datenerhebung – unterschiedliche Methoden der Auswertung, denen unterschiedliche Methoden der *Varianzdekomposition* zugrunde liegen.

Stemmler (1996) systematisiert die möglichen Einflussfaktoren (Person, Variable, Situation/Setting) anhand von drei verschiedenen Dimensionen: (a) „locus of constructs“: Ist das Konstrukt durch Variationen in der Person, der Situation, der Variable oder Kombinationen dieser Aspekte gekennzeichnet? (b) „operatio-

nalisierung“: In welchem Modus (Person, Setting oder Variable) manifestiert sich das Konstrukt? Wo muss die Operationalisierung ansetzen? (c) „range of application“: Auf welche Einheiten beziehen sich mögliche Schlüsse? Bei welchen Modi (Person, Situation oder Variable) sollen sie sich zeigen? Aus diesen drei Dimensionen werden neun verschiedene Messmodelle entwickelt. Sechs dieser Dimensionen entsprechen den Cattellschen Kovariationsschemata (O-, P-, Q-, R-, S- und T-Technik, vgl. Cattell, 1946) und werden als Modelle 1 bis 6 bezeichnet. Drei weitere nehmen explizit die erweiterten Möglichkeiten durch das ambulante Messen auf: Modell 7 bezieht sich auf die Setting-Subjekt-Interaktion, Modell 8 auf die Setting-Variablen-Interaktion und Modell 9 auf die Variablen-Subjekt-Interaktion.

Zwei Beispiele zur Erläuterung: Bei der *R-Technik*, dem am häufigsten verwendeten Messmodell in der Differentiellen Psychologie, werden verschiedene Merkmale (Variablen/Tests) über eine bestimmte Anzahl an Personen korreliert. „Locus of constructs“ ist hier die Person, „operationalization“ bezeichnet die relevanten erhobenen Variablen. Die abgeleiteten Aussagen bleiben auf eine bestimmte Situation beschränkt (range of application). Für die Auswertung nach diesem Modell ist die Between-subject-Varianz relevant, d. h. die Aufdeckung von interindividuellen Unterschieden steht im Mittelpunkt. Im *Modell 7* ist die Interaktion von Setting/Situation und Subjekt das relevante Konstrukt. Personen reagieren in ein und derselben Situation unterschiedlich. Folglich sind die Ergebnisse nur auf gleiche Personen-Situationen-Konstellationen (range of application) transferierbar. Ein Anwendungsbeispiel ist die Messung von Emotionen (Operationalisierungsdimension), bei der davon ausgegangen wird, dass diese nur in einer bestimmten Situation-Personen-Konstellation auftreten.

Die Art des Modells bestimmt dann auch die Auswertung der Daten (vgl. Raudenbush, 1997). Das Messmodell, das sich aus der Fragestellung ergibt, bestimmt die nachfolgende Sampling-Strategie und die Art der Auswertung. Nach Bolger et al. (2003) lassen sich die möglichen Fragestellungen bei der Nutzung des EMA wie folgt systematisieren:

1. *Aggregation über die Zeit*: Was sind typische Personenmerkmale im Hinblick auf die Variabilität über die Zeit? Typische Auswertungsstrategien sind: Bildung von Mittelwerten über die Person und Darstellung der Variabilität; Analyse der Within-subject-Variation; Analyse von Between-subject-Unterschieden in den aggregierten Daten; Analyse beeinflussender Faktoren auf die Between-subject-Unterschiede (vgl. Wittmann, 1985).
2. *Modellierung der Zeitverläufe*: Wie verändern sich die EMA-Daten mit der Zeit im Hinblick auf eine Person und im Hinblick auf Personenunterschiede? Typische Auswertungsstrategien sind: Analyse des Verlaufs bei einer Person; Ana-

lyse von Between-subject-Unterschieden im Hinblick auf den Verlauf; Analyse beeinflussender Faktoren im Hinblick auf Unterschiede in den Verläufen.

3. *Modellierung von Within-subject-Prozessen*: Was sind typische Within-subject-Prozesse und wie unterscheiden sich Personen im Hinblick auf diese Prozesse? Typische Auswertungsstrategien sind: Analyse von durchschnittlichen Within-subject-Prozessen, d. h. Variableninteraktionen bei einer Person; Analyse von Between-subject-Unterschieden im Hinblick auf diese Prozesse; Erklärung von Between-subject-Unterschieden durch weitere Variablen in Hinblick auf die Within-subject-Prozesse.

Die statistische Analyse der Daten zur Beantwortung solcher Fragestellungen stellt eine besondere Herausforderung dar. „Data analysis of MA [= momentary assessment] is particularly challenging because the kind of data collected in EMA studies do not lend themselves to simple traditional analysis ...“ (Shiffman & Stone, 1998, S. 128). Gängige *Auswertungsmöglichkeiten* wie die Aggregation der Daten über Personen, Klassen von Events, Zeiträumen oder Situationen hinweg und die Verwendung einfacher Varianzanalysen, nutzen nicht das Potenzial der Daten (zur Begründung vgl. Schwartz & Stone, 1998). Zur Auswertung sind mixed-effect models (MRCM = multilevel random coefficient modeling) die Methoden der Wahl. Die Darstellungen von Schwartz und Stone (1998), Bolger et al. (2003), Beal und Weiss (2003) und Nezlek (2008) geben eine brauchbare Einführung in die Anwendung der Methode im Hinblick auf die EMA-Daten. Hedeker, Mermelstein und Demirtas (2008) erläutern beispielhaft an einer Studie zum Einfluss von Stimmung auf das Rauchverhalten, mit welchen statistischen Methoden Zufallseffekte auf Ebene der Testpersonen zur Aufklärung von Within-subject-Varianzen modelliert werden sollten. Collins (2006) erläutert, mit welchen Methoden bei längsschnittlichen ambulanten Daten kontinuierliche und diskrete (situative) Veränderungen separiert werden können (vgl. auch Jahng, Wood & Trull, 2008). Eine Alternative zur Zeitreihenanalyse als Analyseinstrument natürlicher Zirkularitäten ist die Anwendung der zirkulären Statistik, die in der Psychologie weitgehend unbekannt ist, aber im Hinblick auf ambulante Daten im Beitrag von Kubiak und Jonas (2007) erläutert wird.

Mithilfe von Mehrebenenmodellen gelingt es, (a) die Probleme möglicher Autokorrelationen und der Heteroskedastizität der Messwiederholungsdaten bei der Parameterschätzung in den Griff zu bekommen, (b) auch mit unvollständigen Datensätzen umzugehen, (c) den zeitlichen Verlauf in Form von Trends, zyklischen Schwankungen und die hierarchischen Abhängigkeiten von Messzeitpunkt und Person zu kontrollieren, (d) die Abhängigkeit und direkten Beeinflussung von EMA-Daten nicht nur durch das Individuum und den Messzeitpunkt, sondern auch durch den sozialen Kontext (Drei-Ebenen-Modelle) zu bestimmen (vgl. Raudenbush, Brennan & Barnett, 1995). Bolger et al.

(2003), Pinheiro und Bates (2000) sowie Singer und Willett (2003) zeigen geeignete Auswertungsalternativen mit gängigen Statistikprogrammen auf.

Mit steigender Anzahl von Einflussfaktoren (Situation, Kontext, Personen etc.) wächst auch die Vielfalt an Auswertungsmöglichkeiten. Es scheint daher ratsam, einen Kanon an notwendigen Informationen und Statistiken festzulegen, wie es beispielsweise schon für die Bewegungsmessung diskutiert wird (Berger et al., 2008) und vor der Datenerhebung Verläufe und mögliche Interaktionen der Parameter zu berücksichtigen (Houtveen & de Geus, 2009).

## 6 Anwendungsbeispiele

Die Sammelbände von Fahrenberg und Myrtek (1996, 2001), Myrtek (2004) sowie Krantz und Baum (1998) liefern einen aktuellen Überblick über die Möglichkeiten des ambulanten Assessments zur Messung physiologischer und psychologischer Daten. Inzwischen liegen auch erste Überblicksarbeiten für den Einsatz bei einzelnen psychologischen Konstrukte, z. B. bei der Bewegungsmessung (de Vries et al., 2006), sowie für spezielle Anwendungsfelder (z. B. in der Arbeits- und Organisationspsychologie: Klumb, Elfering & Herre, 2009, oder der Klinischen Psychologie: Myin-Germeys et al., 2009; Trull & Ebner-Priemer, 2009) vor. Die Arbeit von Hoppmann und Riediger (2009) analysiert beispielsweise die Anwendung in der Längsschnittforschung und benennt hier vier zentrale Einsatzorte: (1) affektiv-motivationale Entwicklung, (2) Einfluss des sozialen Kontextes auf die Entwicklung, (3) altersbezogene Herausforderungen und funktioneller Status im Alter sowie (4) die Messung der kognitiven Entwicklung. Hinsichtlich der Publikationshäufigkeit dominieren in diesem Forschungsfeld medizinische Themen, z. B. zur „white coat hypertension“, zur Bewegungsmessung oder zur Symptomwahrnehmung. Beispielhaft soll in den nachfolgenden Kapiteln die Anwendung des ambulanten Assessments in der Stressforschung (Abschnitt 6.1) und der Schmerzmessung (Abschnitt 6.2) erläutert werden.

### 6.1 Stressforschung

Eine Simulation von Stressoren mit hoher externer Validität ist im Labor aufgrund der vielfältigen Einflussfaktoren im Feld kaum möglich. Mittels EMA gelingt es, das Stresserleben und die damit einhergehenden physiologischen Parameter am Ort der Stressentstehung unmittelbar zu messen. Gleichzeitig ist es möglich, die Situationswahrnehmung, begleitende Bewegungsmuster, die Wahrnehmung der Stressauslöser, Verlaufsmuster, Copingvarianten und die Effizienz des Copings zu erheben (Perrez & Reicherts, 1996). Anwendungen dieser Methoden finden sich bei der Stressmessung in Testsituationen (Zeier, Häseli & Fischer, 2001), in Flugsituationen (Wilson, 2001), während der Tätigkeit von

Fluglotsen (Vogt & Kastner, 2001) und in unterschiedlichen Phasen der Handlungsausführung am Arbeitsplatz (Rau, 2001; Rau & Richter, 1996; Rutledge et al., 2009). Die parallele Erfassung physiologischer Parameter des Stresserlebens ermöglicht weiterhin die Aufdeckung potenzieller Risikofaktoren. So konnten verschiedene Studien zeigen, dass die dynamische Veränderung physiologischer Parameter wie Blutdruck und Herzfrequenz für die gesundheitlichen Folgen des Stresses von großer Bedeutung sind (Steptoe, 2001). Für die Messung physiologischer Parameter im Felde ergibt sich das Problem, dass die reinen physiologischen Folgen von Stress durch körperliche Aktivitäten überlagert sind. Eine Möglichkeit zur Separierung emotionsinduzierter und aktivitätsinduzierter Einflüsse gelingt über die parallele Erfassung von Bewegungsmustern mittels Akzelerometrie. Hierdurch ist es dann auch möglich, Vergleiche zwischen der Interozeption bei körperlicher und bei emotionaler Erregung zu erheben.

Myrtek (2004) konnte beim Einsatz des *Freiburger Monitoring Systems* zeigen, dass psychologische Parameter (Befindlichkeit) eher durch kognitive Schemata mitbestimmt werden als durch die Interozeption körperlicher Veränderungen. Folglich gibt es nur schwache Korrelationen zwischen „physiological workload indicators“ und der Wahrnehmung dieser vermeintlichen Stresszeichen. Dies hat unmittelbare Implikationen für die Einführung von prophylaktischen Screenings.

Die Studien von Myrtek (2004) sowie von Myrtek, Aschenbrenner und Brügger (2005) verdeutlichen, dass die parallele Messung von Emotionen, von Bewegungsmustern und verschiedenen physiologischen Maßen mittels EMA auch geeignet ist, neue differenzialpsychologische Aspekte mit hohem prädiktivem Wert z. B. für das Auftreten kardiovaskulärer Probleme aufzudecken.

## 6.2 Schmerzmessung

Smith und Safer (1993) konnten zeigen, dass die Angaben in herkömmlichen Schmerzfragebögen nur einen geringen Zusammenhang mit den Angaben einer kontinuierlichen Messung mittels EMA aufweisen. Die Mittelung der EMA-Werte klärt nur etwa 50 % der Varianz der Angaben im retrospektiven Schmerzfragebogen auf. Stärkere Zusammenhänge ergeben sich mit dem maximalen Schmerzniveau des zurückliegenden Befragungszeitraums, mit den Schmerzangaben, die sich auf die Schmerzen am Morgen beziehen, und dem schnellsten und kürzesten Anstieg des Schmerzes im Befragungszeitraum (Stone et al., 1999). Andere Untersuchungen zeigen, dass die retrospektive Schmerzeinschätzung vom aktuellen Schmerzerleben überdeckt wird (Jamison, Scrocco & Paris, 1989) und eine kontinuierliche und unmittelbare Messung auch wegen der erhöhten Schmerzaufmerksamkeit für die Entwicklung eines Therapieschemas genauere Aussagen ermöglicht (Peters & Crombez, 2007).

„Electronic diaries“ ermöglichen die Messung des aktuellen Erlebens und die Erfassung besonderer Verlaufsformen des Schmerzes je nach Erkrankung (Godaert, Sorbi, Peters, Dekkers & Geenen, 2001). Durch die parallele Erfassung anderer Variablen können außerdem Faktoren identifiziert werden, die auf die Schmerzwahrnehmung und deren Verlauf Einfluss haben. Sorbi, Honkoop und Godaert (1996) konnten bei der Erfassung über 10 Wochen hinweg mittels sechs Promptings pro Tag zeigen, dass die Anzahl und Art der „daily hassles“ stark mit den Schmerzsymptomen korreliert. In längsschnittlichen ambulanten Erhebungen gelingt es die Zusammenhänge zwischen Schmerz und Bewegung in Abhängigkeit von täglichen Anforderungen zu erheben (van den Berg-Emons, Schasfoort, de Vos, Bussmann & Stam, 2006). Ebenso ist es möglich, den Einfluss von Copingstrategien, Aufmerksamkeitseffekten, der Schmerzneuheit und der Vorhersehbarkeit auf das Schmerzerleben zu separieren (Buck & Morley, 2006). Neben der Angabe der Schmerzstärke (z. B. mit – dynamischen – visuellen Analogskalen) gibt es für die Schmerzmessung auch Softwareprogramme, die den Schmerzcharakter und die Lokalisation des Schmerzes über ein grafisches Display des PDA-Gerätes ermöglichen (Jamison et al., 2004). Schon bei Kindern ab 8 Jahren ist der Einsatz solcher elektronischer Schmerztagebücher möglich und sichert eine höhere Compliance als Papier-Bleistift-Verfahren (Palermo et al., 2004). Ein systematisches Review zum Einsatz des ambulanten Assessments in der Schmerzdiagnostik konnte 62 Studien (Stand 2008) identifizieren (Morren, Dulmen, Ouwerkerk & Bensing, 2008). Die bisherigen Anwendungen bescheinigen dem Verfahren eine hohe Compliance, die durch kurze Fragebögen, eine Einführung in die Bedienung der Geräte, Einsatz eines Alarms zur Erinnerung an die Eingabe und eine finanzielle Entlohnung positiv beeinflusst werden kann. Ein grundsätzliches Problem bei dem Vergleich der Studien ist die seltene Angabe von Antwortraten und „Dropouts“ sowie die Wiederverwendung von Daten in weiteren Publikationen.

## 7 Bewertung des Ansatzes

Die Methoden des ambulanten Assessments können die klassische Labordiagnostik und die Anwendung traditioneller Fragebögen nicht ersetzen. Durch ihre unterschiedlichen Schwerpunkte ergänzen sich beide Strategien. Beispielsweise erscheint es sinnvoll, die Validität klassischer Papier-Bleistift-Verfahren oder Labormethoden mit den Methoden des ambulanten Assessments hinsichtlich testtheoretischer Werte, Antworttendenzen oder Müdigkeitseffekten zu prüfen.

Eine Gegenüberstellung der beiden konzeptionellen Strategien ist auch deshalb nicht sinnvoll, weil der Einsatz von der jeweiligen Fragestellung und dem sich daraus ergebenden Messmodell abhängt. Da es sich um eine Messstrategie und

nicht um ein spezielles Messverfahren handelt, ist eine generalisierte Darstellung von Reliabilitäten und Validitäten unabhängig vom Messverfahren schwer möglich. Buse und Pawlik (1996) weisen in diesem Zusammenhang darauf hin, dass beim ambulanten Messen ein differenziertes Reliabilitätskonzept notwendig ist: Die *lokale Reliabilität* (Buse & Pawlik, 1996) kann durch die parallele Erfassung von Variablen, z. B. von physiologischen Parametern und Selbsteinschätzungen, berechnet werden. Weiterhin ist eine Reliabilitätsberechnung auf der Ebene aggregierter Daten möglich. So können die aggregierten Werte in zwei verschiedenen Zeitfenstern verglichen oder korreliert werden. Perez, Wilhelm, Schoebi und Horner (2001) unterscheiden weiterhin zwischen *situationspezifischer Reliabilität* (Übereinstimmung der Messwerte in ähnlichen Situationen) und der *Interrater-Reliabilität*. Diese kommt dann zum Einsatz, wenn beispielsweise von verschiedenen Personen, die sich in der gleichen Situation befinden, Angaben zur Situationseinschätzung gefordert sind.

Da inzwischen viele für die Messung mittels PDA entwickelte oder adaptierte Verfahren existieren, ist eine empirische Überprüfung der Validität und der instrumentenabhängigen Sensitivität der Messinstrumente möglich.

Im Folgenden werden die in der Literatur häufig thematisierte Äquivalenzproblematik, die Reaktivität, die Akzeptanz/Compliance, die ethischen und rechtlichen Aspekte sowie die Probleme der Messstrategie erläutert.

## 7.1 Äquivalenz

Für Test- und Messverfahren, die sowohl in Papier-Bleistift-Version oder schon als Computerverfahren vorliegen, muss die Äquivalenz einer Adaptation für das ambulante Assessment geprüft werden. Hank und Schwenkmezger (1996) unterscheiden dabei folgende drei Arten von Äquivalenz:

1. *Erfahrungs-Äquivalenz*: „Experiential equivalence focuses on respondents perceptual, emotional, and attitudinal reactions towards the test formats“ (Hank & Schwenkmezger, 1996, S. 86). Diese Art der Äquivalenz hängt mit dem Ausmaß der Reaktivität (vgl. Abschnitt 7.2) zusammen.
2. *Psychometrische Äquivalenz*: Ambulante Messverfahren müssen eine ähnliche Objektivität, Reliabilität und Validität wie die parallelen Fragebögen besitzen und auch hinsichtlich der „Zusatzkriterien“ Normiertheit und Ökonomie (Sidiropoulou, 1997) vergleichbar sein. Buse und Pawlik (1996) fassen die Ergebnisse im Hinblick auf Leistungs- und Persönlichkeitstests wie folgt zusammen: „Results indicate that laboratory-based test results cannot be considered as equivalent to ambulatory assessment data, and vice versa“ (S. 40). Dieses Ergebnis spricht gegen die Äquivalenz, gleichzeitig

aber auch für die Vorteile der ambulanten Messstrategie: Die Varianz der Messwerte wird eben auch durch die Erhebungssituation mitbestimmt.

3. *Stichprobenspezifische Äquivalenz*: Hierbei ist zu klären, ob es typische Stichprobenverzerrungen bei der Messung gibt. Die Äquivalenz ist beispielsweise nicht gegeben, wenn technische Instrumentarien von bestimmten Personen abgelehnt werden. Bolger et al. (2003) sprechen in diesem Zusammenhang von der Gefahr eines „digital divide“, also einer Zweiteilung in diejenigen Personen, die Zugang zu einer eher technisch orientierten Diagnostik wollen und bekommen, und in diejenigen ohne Zugang oder mit Technikablehnung. Zurzeit fehlt es an Studien zu solchen Stichprobeneffekten. Es ist zu vermuten, dass die Akzeptanz bei Personengruppen, die entsprechende Geräte schon nutzen, positiver ausfällt als bei Techniknovizen.

## 7.2 Reaktivität

Während es bei der Frage der Äquivalenz um die Ähnlichkeit zweier Messmethoden oder Messstrategien geht, bezeichnet der Begriff „Reaktivität“ die Verzerrung der Messung durch die Eigenheiten der Methode, das heißt, die Messung selbst klärt Varianz in den Daten auf. Eine hohe spezifische Reaktivität einer Erhebungsmethode bedroht damit letztlich die Äquivalenz.

Die Reaktivität des ambulanten Assessments lässt sich mittels experimenteller Forschung messen. Grundsätzlich ist die Anzahl an bisherigen Studien zur Prüfung der Reaktivität eher gering. In der Untersuchung von Rowan et al. (2007) wurden zur Messung der Reaktivität bei der Raucherentwöhnung beispielsweise die Anzahl der abstinenten Personen mit und ohne ambulante Messung sowie die Angaben zur Stimmung, Konzentration, Selbstwirksamkeit etc. verglichen. Zur Sicherung der Angaben dienten biochemische Messungen, die Auskunft über die Rauchabstinenz gaben. Die Ergebnisse legen nahe, dass durch den Einsatz des ambulanten Assessments Reaktivitätseffekte möglich sind. Beispielsweise ist bei dauerhaften Kontrollen mittels ambulanter Messung der Anteil der abstinenten Personen höher. Solche Auswirkungen müssen beim Einsatz des ambulanten Assessments beachtet werden.

Unter gleichen messtechnischen Voraussetzungen ist die Reaktivität im Labor sicherlich als höher zu bewerten, da die besonderen Räumlichkeiten und die Anwesenheit eines Testleiters die Messergebnisse beeinflussen können. Das Beispiel des „office hypertension“ (Pickering et al., 1985; vgl. Abschnitt 3) verdeutlicht dies.

Beim ambulanten Messen kann sich eine Verhaltensänderung dadurch ergeben, dass die immer wiederkehrende Aufforderung, auf das Verhalten oder die

Stimmung zu achten, zur Veränderung der Selbstaufmerksamkeit führt (Kazdin, 1974). Ebenso ist durch die wiederholte Messung erst die Bildung des Konstrukts „Veränderung“ bei der Testperson möglich (Bolger et al., 2003). Diese Reaktivitätsursache kann aber auch im Sinne einer Validitätsverbesserung interpretiert werden.

Die Reaktivität durch die erhöhte Selbstaufmerksamkeit hängt auch von der Möglichkeit ab, Einblick in bisherige Datenerhebung zu bekommen. Prüfungen von Hank und Schwenkmezger (1996) mit dem Daily Behavior and Experience Inventory (BEI) zeigen, dass die Erfahrungsvariabilität hoch ist, wenn es möglich ist, die Daten nochmals einzusehen, wie dies zum Beispiel bei klassischen Papier-Bleistift-Tagebüchern der Fall ist. Die Eingabe in ein PDA-Gerät ohne die Möglichkeit, die vorherigen Eingaben nochmals zu sehen, reduziert den Einfluss solcher Erfahrungsvariabilität und damit auch die Reaktivität.

Reaktivität kann sich weiterhin auch durch Verhaltens- und Stimmungsänderungen ergeben, die direkt mit der Dateneingabe in Verbindung stehen. Es ist einleuchtend, dass ein PDA, der in regelmäßigen Abständen die Eingabe von Daten (z. B. zur Stimmung) verlangt, mit der Zeit Einfluss auf die Stimmung nimmt (Tseng, Tiplady, McLeod & Wright, 1998). Schon das Ertönen des Promptings kann dann als Folge einer Konditionierung letztlich Stimmungsschwankungen auslösen, bevor die eigentliche Abfrage beginnt. Unterschiedliche mobile Geräte (Handy, PDA, IVT etc.) können dabei zu unterschiedlichen Einflüssen auf die Stimmung führen und daher das Messergebnis beeinflussen (Isomursu, Tähti & Väinämö, 2007). Insbesondere wenn der kognitive und motorische Aufwand für die Bedienung solcher Geräte hoch ist und die eigentliche Beantwortung durch einen schlechten Bedienkomfort überlagert ist, kann es zu entsprechenden Verzerrungen kommen.

Zwischen dem Versuch, kritische Erlebensaspekte in ihrem Verlauf exakt zu messen (z. B. starker Schmerz), und der Gefahr, Personen dadurch zu stark zu belasten, muss eine Balance gefunden werden. Ebenso sind konfligierende Instruktionen für zeitgleiche Messungen ein Grund für Reaktivitätseffekte. Beispielsweise sind die parallele Erfassung des Blutdrucks und das gleichzeitige Ausfüllen des Fragebogens kaum möglich, weil für das Blutdruckmessen mit klassischen Manschetten in der Ellenbeuge eine Armstreckung notwendig ist. Hier sind technische Weiterentwicklungen gefragt, die dazu führen, dass die Erhebung noch weniger aufdringlich ist und die Instruktion genauer umgesetzt werden kann (vgl. Seibt, Berndt, Knöpfel & Scheuch, 2001).

Buse und Pawlik (1996) zeigen, dass bei Erhebungen über längere Zeiträume die Reaktivitätseffekte besonders bei *den* Aspekten des Verhaltens zunehmen, die von der Person selbst negativ bewertet werden (z. B. Essattacken, Drogen-

konsum). Beim Sampling mit fixierten Intervallen wird gezielt Verhalten gemieden, das während des vorhersehbaren Messzeitpunktes auftreten könnte. All diese Beispiele zeigen, dass die Art der Stichprobenbildung und die Häufigkeit der Messung einen bedeutsamen Einfluss auf die Reaktivität haben können.

### 7.3 Akzeptanz und Compliance

Die Notwendigkeit mehrfacher Dateneingaben durch die Testperson kann im Vergleich zum einmaligen Ausfüllen eines Fragebogens die Akzeptanz und die Compliance bedrohen. Akzeptanz bezieht sich auf die Einstellung der Testperson bei der ambulanten Messung mitzuwirken. Diese *kann* die tatsächlich gezeigte Mitwirkung bei der Datenerhebung beeinflussen. Die Compliance wird aber auch durch die Teilnahmemotivation, eine mögliche Honorierung oder einen großen persönlichen Nutzen der Messung beeinflusst (Simpson et al., 2005). Eine geringe Akzeptanz des Verfahrens kann bei entsprechendem subjektiven Nutzen also dennoch mit einer hohen Compliance einhergehen. Besonders für physiologische Messungen ist es wichtig, dass die Messvorrichtung als angenehm empfunden wird. Schlecht klebende Elektroden, eine geringe Passform oder bewegungseinschränkende Vorrichtungen reduzieren die Akzeptanz.

Die Akzeptanz lässt sich durch qualitative und quantitative Nachbefragungen klären. In der Regel ist die Akzeptanz von computerisierten Eingabeformen beim ambulanten Assessment höher als bei Papier-Bleistift-Bögen, jedoch kommt es bei klinischen Stichproben häufiger zu unvollständigen Angaben und möglicherweise auch technischen Ausfällen, die die Datensammlung bedrohen (Gallagher et al., 2008). Fahrenberg et al. (2007) resümieren im Hinblick auf die Akzeptanz, „dass die durchschnittliche Akzeptanz für die Methodik bei den Untersuchten weitaus größer ist als die Bereitschaft in der Psychologie, diese innovative Methode in Forschung und Praxis zu nutzen“ (S. 18).

Die *Compliance* lässt sich über die „response rate“ bestimmen: Auf wie viele Promptings erfolgte durch die Testperson eine Eingabe in das Eingabegerät? Es ist zu bedenken, dass eine fehlende Dateneingabe die Repräsentativität der Daten und damit letztlich die Reliabilität und Validität bedroht; insbesondere dann, wenn es sich um nicht zufällige Non-Compliance handelt. Die Bereitschaft, auf ein Prompting zu reagieren, liegt in der Regel zwischen 80 und 90 % (vgl. Beal & Weiss, 2003; Hank & Schwenkmezger, 1996). Nur 20 bis 30 % der Personen, die ein solches Instrumentarium benutzt haben, bewerten es im nachhinein als störend (vgl. Hank & Schwenkmezger, 1996). Selbst bei schwer erkrankten Patientengruppen, beispielsweise Patienten mit Stammzell-

transplantation, ist die Messung mittels PDA durch eine hohe Akzeptanz und hohe response rate gekennzeichnet (Hacker & Ferrans, 2007). Die Verwendung hochwertiger Technik sichert besonders bei technikbegeisterten Personen die Compliance. Im Vergleich zu „nicht momentanen“ Fragebögen ist es für die Testpersonen beim EMA leichter, Angaben zur aktuellen Befindlichkeit zu machen, da keine Reflexionsprozesse über die zurückliegenden Zeiträume notwendig sind. Grundsätzliche Einstellungen gegenüber den technischen Geräten können die Akzeptanz beeinflussen, aber leicht durch vorherige Messungen bestimmt werden. Zur Messung der Einstellung gegenüber Handheldgeräten sind beispielsweise Fragebögen entwickelt worden, die im Unterricht und im Trainingsbereich Verwendung findet (z. B. zur Messung der „Attitude Toward Handheld Devices“; Van't Hooft, Díaz & Swan, 2004).

Die Akzeptanz und Compliance werden weiterhin durch den Aufwand bei der Dateneingabe beeinflusst. Die Gefahr besteht eher in einer Belastung durch die Summe vieler Erhebungszeitpunkte als durch eine einzelne Dateneingabe. Das muss in entsprechenden Vortests eruiert werden. Wird die Person mit einer Vielzahl von Fragen konfrontiert, dann sinkt schnell die Teilnahmemotivation. Analysen von „Missing data“-Häufigkeiten zeigen meist einen U-förmigen Verlauf: Anfängliche Fehleingaben oder fehlende Eingaben sind durch unzureichendes Training und damit Unsicherheiten beim Handling zu begründen, spätere „missings“ sind durch eine abnehmende Motivation der Testpersonen verursacht (Flechter et al., 2003).

Neben dieser sogenannten „signal compliance“ gibt es auch eine „setting compliance“. Hier geht es um die Frage, in welchen Situationen Personen eher bereit sind, Angaben zu machen. Eine Eingabeaufforderung beim Theaterbesuch oder im Schwimmbad bleibt eher unbeantwortet als eine in der Arbeitswelt oder in der häuslichen Umgebung. Bei der Anwendung des EAR (Electronically Activated Recorder), der automatisiert Geräusche und Sprache aufzeichnet, zeigt sich eine geringere Compliance, d. h. ein selteneres Tragen der Aufzeichnungsgeräte, wenn sie am Wochenende getragen werden sollen (Mehl & Holleran, 2007). Solche Aspekte müssen bei der Konzeption der Studie und beim Sampling beachtet werden, da sonst die Repräsentativität der Daten bedroht ist. Eine mangelnde setting compliance kann sich durch fehlende Angaben in bestimmten Situationen oder durch verzögerte Angaben (das sog. „hoarding“) bemerkbar machen. Die Testpersonen verschieben dabei die Eingabe auf spätere Gelegenheiten, in denen eine ungestörte oder unauffällige Eingabe möglich ist. Dies setzt natürlich die Möglichkeit voraus, die Angaben verspätet zu machen. Durch enge Zeitfenster zwischen zwei Eingabeaufforderungen kann dieses „hoarding“ verhindert werden; allerdings geht dies mit einer steigenden „Missing-data“-Quote einher. Erfolgen häufige Erinnerungen an die Dateneingabe und besteht keine Möglichkeit, die Aufgabaufforderung abzubrechen, kann es zur

Eingabe von falschen Daten kommen („fake compliance“). Umfang und Art der Compliance hängen daher stark von der Erhebungsstrategie ab. Folgende Maßnahmen erhöhen die Compliance:

- a) Es muss ein Training mit der Erhebungstechnologie stattfinden. Dies gilt insbesondere bei der Verwendung von PDA-Geräten und komplexen Items.
- b) Den Testpersonen sollte die Möglichkeit gegeben werden, die Beantwortung der Items auf einen späteren Zeitpunkt zu verschieben oder in begrenztem Umfang Angaben zu verweigern, also „Kontrolle über die Kontrolle“ zu behalten. Zusammen mit der Möglichkeit, Eingaben zu verschieben, kann schon eine elektronische Erfragung von Gründen stattfinden. Weiterhin sind im *Post-monitoring-Interview* die Gründe für die Compliance zu erfragen (Stone & Shiffman, 2002). Für die Langzeit-EKG-Messung ist beispielsweise ein spezieller Post-monitoring-Interviewleitfaden entwickelt worden (Kessler, Chakko & Kessler, 1994).
- c) Hilfreich sind auch automatisierte Erinnerungen an nicht erfolgte Einträge, die über das Gerät, durch das Telefon oder andere Signalgeber erfolgen können. Umfang und Art der Erinnerungen sind jedoch sorgfältig zu wählen, um bei den Testpersonen nicht Reaktanz auszulösen. Bei Ho und Intille (2005) wird versucht, Maße zu finden, die die Bereitschaft der Testpersonen widerspiegeln, auf ein Prompting zu reagieren. Mithilfe akzelerometrischer Verfahren ist es möglich, die „receptivity to interruptions“ zu bestimmen und damit Informationen über die weitere Bereitschaft zur Dateneingabe zu erhalten. Die Identifikation von Phasen mit hoher Handlungsunterbrechungsneigung erlaubt die Identifikation von Zeiträumen, in denen die Compliance z. B. zur Dateneingabe in ein elektronisches Tagebuch besonders groß ist (Ho & Intille, 2005). Hier wird also mit diagnostischen Mitteln geprüft, wann die Bereitschaft auf das Prompting zu reagieren und damit die Compliance am höchsten ist.
- d) Die Bereitschaft zur Dateneingabe hängt auch von der Bedienfreundlichkeit des Messinstrumentes ab. Die Usability der Geräte und die Compliance sollten in einem Pretest geprüft werden. Solche Vortests können Schwachstellen in der Bedienung offenlegen und beispielsweise Verbesserungen der Itempräsentation auf den PDA-Geräten nach sich ziehen (Stinson et al., 2006).
- e) Die Geräte sollten möglichst harmonisch in das tägliche Leben integriert werden, d. h. transportabel sein und eine unauffällige Dateneingabe ermöglichen. Hufford et al. (2001) nennen diese Möglichkeit „livability functions“. Dies ist besonders bei der Verwendung von Handys zur Dateneingabe mittels IVR gegeben. Im Konzept der „ambient intelligence“ (vgl. Abschnitt 8.3) ist die unauffällige Messung definierender Bestandteil.
- f) Bei der Darbietung von Itemlisten sollte die Abfolge und Auswahl der Fragen adaptiv erfolgen. So sind Kürzungen des Fragebogens oder Einschränkungen

kungen auf wesentliche Items in Abhängigkeit von der Compliance möglich. Ebenso ist auch eine Auswahl relevanter Items aufgrund adaptiver Schwierigkeitsanpassungen möglich.

- g) Es sollte in der Instruktion an die soziale Verantwortlichkeit („accountability“) appelliert werden, indem zum einen die Wichtigkeit einer exakten Dateneingabe für die Diagnostik betont wird und zum anderen darüber informiert wird, dass die Messung der Compliance ein Teil der Erhebung ist. „If subjects know that they will be held accountable for their data, they tend to be more vigilant with the diary protocol“ (Hufford et al., 2001, S. 86).
- h) Eine direkte Rückmeldung der Ergebnisse und der Auswertung sowie ein Feedback über die Eingabequote fördern ebenfalls die Compliance. Dies gilt besonders dann, wenn aus der Auswertung Handlungsempfehlungen abgeleitet werden, wie dies bei der Telemedizin der Fall ist.

#### 7.4 Ethische und rechtliche Aspekte

Im Hinblick auf forschungsethische Aspekte schreibt Fahrenberg (1996): „Ambulatory assessment may violate privacy more easily than other methods“ (S. 14). Diese Gefahr begründet sich daraus, dass hierbei an die Stelle einer einmaligen Testung längere Erhebungszeiträume mit mehrfachen Messungen außerhalb der Laborsituation treten können. Die Testperson ist damit gezwungen, prospektiv ihr Einverständnis zur Datenerhebung in Situationen zu geben, die sie beim Training oder zu Beginn der Instruktion nicht abschätzen kann.

Daher ist es ratsam, die Personen vorab, über Abbruchmöglichkeit und die Verwertung der bis dahin angefallenen Daten zu informieren. Dieses Vorgehen stellt sich jedoch bei physiologischen Messungen als schwer dar, denn hier bedeutet „Abbruch“ nicht die Verweigerung einer aktiven Dateneingabe, sondern die Beendigung der passiven Messung. Ein Abbruch ist damit für die Testperson nicht so leicht möglich.

Im Rahmen eines Post-monitoring-Interviews sollte daher das Einverständnis über die Nutzung der Daten thematisiert werden. Diese nachträgliche Zustimmung muss auch deshalb eingeholt werden, weil die Messinstrumente so unaufdringlich sind, dass das Bewusstsein für die Messung im Feld oft fehlt. Diese Gefahr ist durch die weitere Minimalisierung von Messgeräten und die Möglichkeit einer ubiquitären Messung ein wachsendes Problem. Für eine umfassendere ethische Auseinandersetzung vgl. Schuler (1982) sowie Bersoff und Bersoff (2000).

Mit den ethischen Aspekten sind auch datenrechtliche Fragen verbunden. Der Zugang zu gespeicherten Daten ist durch das Bundesdatenschutzgesetz gere-

gelt. Demnach ist die Zustimmung zur Speicherung einzuholen und der Zugriff auf die Daten auf einen eingeschränkten Nutzerkreis einzugrenzen, für den die Person vorab das Einverständnis gegeben hat. Die Sicherung der Daten ist in der Labordiagnostik leicht kontrollierbar, bei der ambulanten Messung liegt der Schutz der Daten aber bei der Testperson selbst. Wer ist verantwortlich, wenn PDAs mit sensiblen Daten aus dem Besitz der Testperson entwendet werden? Auf solche Möglichkeiten ist die Person zu Beginn der Erhebung hinzuweisen. Werden Daten telemetrisch übertragen, sind entsprechende Sicherheitsprotokolle zu wählen, die eine verschlüsselte Übermittlung garantieren.

#### 7.5 Probleme und Nachteile des ambulanten Assessments

Während der Einsatz von Computern in der Diagnostik stetig wächst, scheint dies für die ambulanten Messmethoden nicht der Fall zu sein (Richard & Bobicz, 2003). Vergleicht man die Einsatzhäufigkeit in der Medizin mit der in der Psychologie, dann fällt auf, dass hier eine große Lücke klafft. So bringen die Begriffe ambulatory assessment und ambulatory monitoring der Datenbank MedLine für die Jahre 1950 bis 2005 insgesamt 6045 Einträge, während man in PsycINFO 188 und in PSYINDEX – ergänzt um die passenden deutschen Begriffe – gerade einmal 91 Einträge findet. Fahrenberg et al. sehen daher innerhalb der Psychologie noch ein erhebliches Nutzungspotenzial: „Die verfügbare Technik ist weiter entwickelt als sie in der Forschung und professionellen Praxis der Psychologie Anwendung findet“ (Fahrenberg et al., 2007, S. 14).

Als Gründe für den geringen Einsatz werden von Fahrenberg, Foerster und Müller (1996) und Fahrenberg (2001) folgende Argumente angeführt:

1. Die geringe Verwendung technischer Geräte ist durch eine negative Einstellung gegenüber technischen Innovationen bedingt. Dieses *Image* wurde durch die Probleme der Technikbedienung in den 80er und 90er Jahren geprägt. Damals waren intensive Trainings der Testleiter und der Testpersonen für die Techniknutzung notwendig. Die Bedienung war problematisch und für die Nutzung war man auf die Mithilfe von Informatikern angewiesen. Diese technophobe Einstellung behindert bis heute den Einsatz in Forschung und Praxis, obwohl die Bedienbarkeit von Hard- und Software nicht mehr mit dem Stand in den 90er Jahren vergleichbar ist.
2. Ein Kennzeichen des EMA ist die Beachtung der Setting-Einflüsse. Dies weicht von dem in der Psychologie favorisierten hypothetiko-deduktiven Vorgehen (vgl. Erdfelder & Funke, 2004) ab, bei dem man sich bemüht, im Experiment den Einfluss von Kontextfaktoren zu minimieren, um die interne Validität zu sichern. Der *methodische Zugang* ist daher ungewohnt.

3. Ein weiteres Argument sind die hohen *Kosten* für die Hard- und Software. Hier hat sich in den letzten Jahren jedoch eine deutliche Kostenreduktion ergeben. Gute Softwareprogramme sind kostenlos verfügbar (Le & Hat, 2006) und Smartphones mit Farbdisplay und großen Speichermöglichkeiten sind für unter 100 Euro erhältlich. Die Anschaffungskosten beim ambulanten Messen müssen im Übrigen auch gegenüber der Einrichtung von teureren Laborräumen oder der Anschaffung von Computertestsystemen abgewogen werden. Folgekosten entstehen aber durch den notwendigen technischen Support für die Wartung und die Suche nach geräte- oder programminduzierten Fehlerquellen.

Ergänzend dazu sind folgende Probleme zu nennen:

1. Die *technischen Möglichkeiten* mobiler Geräte sind gegenüber stationären Computerverfahren noch sehr eingeschränkt. So sind die Displays sehr klein und die Eingabe der Daten nicht so bequem möglich wie über eine Standard-Tastatur.
2. Insbesondere bei physiologischen Messungen fallen *große Datenmengen* an, die zudem noch mit vielen *Störsignalen* verbunden sind. Hier müssen Filtermethoden entwickelt werden, die entsprechendes physikalisches Know-How und meist auch Programmierkenntnisse voraussetzen. Die Entwicklungsarbeit und auch die Bildung von sinnvollen Auswertungsrationalen für physiologische Daten sind letztlich nur interdisziplinär sinnvoll zu lösen. Zur Entwicklung von Auswertungsalgorithmen gibt es erste Empfehlungen von Schwartz und Stone (1998).
3. Der Entwicklung von entsprechenden Verfahren im universitären Kontext muss auch eine *Nutzung in der Praxis* folgen. Es muss also schon durch das Studium sichergestellt werden, dass das entsprechende Wissen in die Praxis Einzug hält. Eine Darstellung der technischen Möglichkeiten beim ambulanten Assessment in der Lehre ist daher zu empfehlen. Entwicklungsarbeiten müssen von den Kriterien Verfügbarkeit, Zugänglichkeit und Erschwinglichkeit für die Praxis geleitet sein, damit diese Messstrategie nicht ein Refugium von Forschungseinrichtungen bleibt.
4. Die Erhebung im Felde macht besondere Sampling-Strategien notwendig. Für spezielle psychologische Variablen (z. B. Müdigkeit) fehlt es noch an *Konventionen* über die Art des Samplings, d. h. an klaren Regelungen zur Dauer der Erhebung, zum Zeitfenster und der Häufigkeit der Promptings. Während es z. B. in der medizinischen Diagnostik klare Regelungen über die Messprozedur für Blutdruck und Langzeit-EKGs gibt (z. B. Middeke et al., 1992), fehlt Entsprechendes in der Psychologie. Umso wichtiger ist eine detaillierte Darstellung der verwendeten Technik und der Stichprobe in Publikationen. Die „Reporting Guidelines“ von Stone und Shiffmann (2002) sowie Berger et al. (2008) sind hierfür eine Hilfe.

5. Wie im Kapitel 5 dargestellt wurde, setzt die Auswertung von EMA-Daten hohe *methodische Kompetenzen* voraus. Einfache Aggregationen von Daten auf dem Niveau von Mittelwerten nutzen nicht die vielfältigen Möglichkeiten einer modellgeleiteten Varianzdekomposition. Zur adäquaten Auswertung sind daher statistische Methoden für Mehrebenen-Modelle notwendig.

## 8 Zukünftige Entwicklung des ambulanten Messens

Ein Blick auf technische Trends lässt erahnen, wie sich die Diagnostik in den nächsten Jahren verändern wird. Zwei Entwicklungen sind dabei von besonderer Bedeutung: Die Entwicklung von RFID-Chips und die erweiterten Möglichkeiten zur Lokalisierung und Aufzeichnung von Bewegungsdaten. Die dabei erhobenen Daten stellen aufgrund der vielen Messzeitpunkte und der Erhebung *in situ* eine Herausforderung für die Datenauswertung dar. Sie bieten aber auch eine Chance für die Weiterentwicklungen der diagnostischen und statistischen Grundlagenforschung. Die traditionelle Rolle des Diagnostikers als Testanwender und -auswerter wird mit den sogenannten „smart technologies“ und deren Nutzung in „Smart-home-“ oder „Ambient-intelligence“-Konzepten aufgegeben. Diese Entwicklungen ermöglichen eine automatisierten Auswertung, *Rückmeldung* und auch *Rückwirkung*.

### 8.1 Erweiterte technische Möglichkeiten in der Diagnostik

Ein Rückblick auf die Geschichte der Diagnostik verdeutlicht, dass es bestimmte Trends in der Wahl der Erhebungsmethoden gab (z. B. Schwarb, 1996, für die Entwicklung der Diagnostik in der Personalauswahl). Technologische Entwicklungen wurden dabei mit einer zeitlichen Verzögerung in der Psychologie zur Kenntnis genommen und bestimmten dann einen diagnostischen Trend. Die Psychologie reagierte eher auf technische Entwicklungen, als diese aktiv mitzugestalten. Dies führte auch dazu, dass oft die Methode die Inhalte bestimmte und nicht umgekehrt.

Betrachtet man die derzeitigen technischen Entwicklungen, dann ist schon jetzt absehbar, dass sich neue diagnostische Möglichkeiten des ambulanten Assessments ergeben werden, an deren Entwicklung Psychologen aktiv mitwirken sollten. Fachspezifische Entwicklungen für diagnostische Zwecke machen eine interdisziplinäre Zusammenarbeit mit Ingenieuren, Computerwissenschaftlern und Informatikern notwendig.

Welche *Trends* sind nun in den nächsten Jahren zu erwarten und welche Auswirkungen werden dies auf die Diagnostik haben? Die zunehmende *Miniaturi-*

sierung von Messvorrichtungen wird dazu führen, dass die Messung unauffälliger und unaufdringlicher wird. Die Geräte werden „smart“, d. h. ohne Beeinträchtigung der Person in der Lage sein, Messung weitgehend automatisiert vorzunehmen. Das „LifeShirt“ der Firma VivoMetrics (Wilhelm et al., 2003) ist eine solche „smarte“ Technologie. Es handelt sich um eine Vorrichtung zur Erhebung physiologischer Daten, die wie ein T-Shirt getragen und gewaschen werden kann. Nur ein Datenrekorder erinnert daran, dass hier eine Messung stattfindet. Mit dem Konzept des „ambient home“ (vgl. Abschnitt 8.3) wird die Idee einer unauffälligen Messung im häuslichen Bereich besonders weitreichend umgesetzt. Hier finden verschiedene Datenerhebungen statt, die automatisiert zu Anpassungen und Wirkungen führen. Die unauffällige und in den Alltag integrierte Technologie erhöht die Akzeptanz und Compliance, birgt aber auch die Gefahr, dass die Reflexion der Testpersonen über die laufende Messung schwindet. Eine radikale Form der Minimalisierung stellt die *Nanodiagnostik* dar (Ferrante, 2005). Diese in der Medizin entwickelten Verfahren ermöglichen es, im Inneren des Körpers relevante Daten zu sammeln und diese drahtlos an externe Auswertungsgeräte zu senden. Die Minimalisierung der dabei verwendeten Sensoren erreicht inzwischen die Größe einer Tablette. Eine Messung ist so völlig „unobstrusiv“ möglich.

Es wird in Zukunft eine *verstärkte Anpassung der Messinstrumente an die Fragestellungen der Diagnostik* geben. Wir befinden uns derzeit noch in der Phase, in der versucht wird, die Gestaltung der Items an die Geräte anzupassen. So müssen bei der Messung mittels PDA die Items kurz sein, damit sie auf dem Display lesbar sind. Umfangreiche Datenspeicherungen, z. B. von digitalisierten Sprachmitteilungen, sind derzeit auf den mobilen Geräten nur begrenzt möglich. Mit den wachsenden technischen Möglichkeiten werden diese Restriktionen sicherlich verschwinden.

Mit der *Verbesserung der Bedienbarkeit* der Geräte wird die Diagnostik nicht nur vermehrt aus dem Labor ins Feld verlagert, sondern auch vermehrt in die Verantwortung der Testperson übergehen. Mit der Möglichkeit von Online-Tests ist schon ein erster Schritt in diese Richtung gemacht. Denkbar ist z. B., dass zukünftig auch relevante Laborparameter (z. B. der Kortisolspiegel) selbstständig bestimmt und dann telemetrisch an die Testleiter weitergeleitet werden.

Eng damit verbunden ist eine zunehmende *Freizeitorientierung* von ambulanten Messmethoden. Schon heute haben Leistung- und Persönlichkeitstests Einzug in verschiedene Fernsehshows genommen. Testungen an und mit den Zuschauern sind durch „Televoting-Systeme“ möglich, bei denen Items in einer Fernsehsendung präsentiert werden und die Antworten mittels Telefon übertragen werden. In der häuslichen Umgebung werden zukünftig mit Spielekonsolen

Gruppen- und Einzeltestungen möglich sein. Durch die neuen Möglichkeiten einer bewegungsgesteuerten Eingabe (z. B. mittels Wii-Konsole; Gamberini, Cardullo, Seraglia & Bordin, 2010) werden beispielsweise auch standardisierte Bewegungsmessungen möglich sein. Wir steuern mit diesen vereinfachten Messmöglichkeiten und der automatisierte Auswertung auf eine *Boulevardisierung der Diagnostik* zu. Hier ist es wichtig, dass die Psychologie diese Entwicklung im Auge behält und die Qualität entsprechender Verfahren sichert bzw. sich auch in der Öffentlichkeit von ungeeigneten Messmethoden distanziert.

Durch *leistungsfähigere Prozessoren, größere Speichermöglichkeiten* und bessere bildliche Darstellungsformate wird die Darbietung komplexer Stimuli erleichtert („media-rich assessments“, vgl. Jones & Dages, 2003). Videotestungen mit hervorragender Bildqualität werden zukünftig auf Smartphones auch im Feld möglich sein.

*Sinkende Preise* bei der Hard- und Software werden die Nutzung auch für die psychologische Praxis attraktiver machen. Damit wird die Phase von Pilotstudien sicherlich bald überwunden sein. Schon jetzt gibt es für physiologische Messgeräte und entsprechende Softwareprodukte einen großen Markt, der sich auf die Preisentwicklung für die Endnutzer günstig auswirkt. Neben den PDA-Geräten und Smartphones werden auch mobile Rechner kleiner und kostengünstiger. Mit der Zunahme von sogenannten Subnotebooks (Netbooks), UltraMobile PCs und Tablet PC werden noch anwenderfreundlichere Testformate bei gleicher Mobilität der Messgeräte möglich sein.

Neue *Programmiersprachen und Betriebssysteme* werden die Entwicklung von Softwareprodukten für ambulante Messinstrumente verbessern und beschleunigen. Die Entwicklung technischer Diagnostika wird damit nicht mehr nur Informatikern vorbehalten sein, sondern auch vermehrt von Psychologen selbst übernommen werden können. Die mit der Softwareverbesserung einhergehende Usability-Steigerung wird die Compliance und Akzeptanz der Testpersonen erhöhen. Weiterhin wird die zunehmende *Verbesserung der Datenübertragung* durch WLAN, Bluetooth oder UMTS auch für das ambulante Assessment neue Möglichkeiten bringen. Die neuen Sendetechniken bei Mobiltelefonen werden den Anteil an telemetrischen Verfahren erhöhen und den Austausch von Daten über größere Distanzen hinweg fördern. Ambulantes Assessment und Telemetrie werden sich immer mehr angleichen (Leopold, 2001). Mit den erweiterten Möglichkeiten der Datenübertragung wird auch die *Vernetzung der technischen Instrumentarien* (PDA, Handy, Computer, RFID-Chips, Satellitendaten, „smarte“ Sensoren etc.) zunehmen. Hierdurch ist ein schneller Vergleich von Daten zur Aufdeckung von Kontingenzen und Korrelationen möglich. Das dichter werdende Netz an „Diagnosestellen“ und Verbindungen dieser Stellen wird dann zu einer „neural net driven scoring platform“ (Jones & Dages, 2003)

anwachsen. Zugriffe auf diese Plattform sind von verschiedenen Orten der Welt aus möglich, ebenso kontinuierliche Einspeisungen neuer personenbezogener Daten. Durch die Möglichkeit, Testgeräte verschiedener Personen und andere diagnostische Stellen (z. B. Sensoren im ambient home) schnell durch mobile Ad-hoc-Netzwerkbildungen (sogenannte MANets oder Meshnetze) zu verbinden, werden gezieltere und situative Normierungen entstehen. Beispielsweise ist es heute schon mithilfe vernetzter PDA-Geräte möglich, den Lernfortschritt von Schülern in Relation zur Klassenleistung zu messen. Zusätzlich können vom Lehrer schnell neue Testitems und Lernaufgaben an die Klasse oder Klassenstufe mittels drahtloser Technologien verteilt werden (Roschelle, Penuell & Yarnall, 2005).

Die umfangreiche Vernetzung wird zu großen *Datenbanken* führen, die weltumspannende systematische Normierungstichproben, ein Benchmarking verschiedener Testsysteme, die Qualitätssteigerung von Verfahren und eine automatisierte Aktualisierung der Software, der Tests und einzelner Items ermöglichen. Als weitreichende Form einer umfassenden personenbezogenen Datensammlung sind die Bestrebungen des „MyLifeBits-Projekt“ von Microsoft zu nennen (Gemmell, Bell & Lueder, 2006). Der Wert dieser komplexen Datensammlung und -speicherung liegt in der Nutzung entsprechender Daten für die Bewältigung des Alltags. Mit entsprechenden Modellen und Auswertungen kann es beispielsweise gelingen, Computersysteme adaptiv zu gestalten. So lässt sich aus der Analyse der gewählten Internetseite und der Lesedauer auf die Lesefertigkeit schließen und adaptiv werden einfachere Seiten gezeigt („personal simplified web“). Die besondere Herausforderung dieses Projektes ist nicht die Datenerhebung, sondern die Speicherung und sinnvolle Strukturierung der Daten. „This tsunami of data presents numerous challenges to computer science, including: how to physically store such „digital memories“ over decades“ (Fitzgibbon & Reiter, 2003, S. 1). Hierzu sollen entsprechende Vernetzungen geschaffen werden, die es beispielsweise ermöglichen, anhand von Fotos Verbindung zu allen „Lebensdaten“ („digital memories“) der dort fotografierten Personen, Gegenstände und Lokalitäten zu erhalten. Die Speicherung dieses „Lebensgedächtnisses“ wird durch die Vernetzung verschiedener Systeme ermöglicht: „The challenge depends on a paradigm shift: from the concept of a ‚computer‘ for life, to a ‚memory‘ for life: an information repository which is conceptually separate from the computers which manage that repository. In turn, we must stop thinking of information in terms of disjoint data types (images, audio, recording, text files, web pages), and instead think holistically of information as giving different perspectives on people, events, and the world.“ (Fitzgibbon & Reiter, 2003, S. 6). Den Verlockungen, hier entstehe eine riesige Datenbank mit Längsschnittdaten über Bewegungsdaten, Daten über Kontakte, Wechsel von Wohnorten und Arbeitsstätten etc., steht das Problem gegenüber, die gigantischen Daten zu strukturieren (Stichwort „data

mining“) und ausgehend von theoretischen Erwägungen zu nutzen. Die Mühen, die es macht, gleichzeitige Messungen („congruent assessment“) bei physiologischen Mehrkanal-Rekordern sinnvoll zu reduzieren, ist vermutlich nur ein kleiner Ausschnitt der Anforderungen, die auf die Statistiker zukommen werden, wenn es um so umfassende Datensammlungen wie beim MyLifeBits-Projekt geht. Aufgrund der besonderen Komplexität und Vernetztheit der Strukturen wird die Erforschung des virtuellen „memory of life“ ebenso wissenschaftliches Know-How, Zeit und Anstrengung erfordern, wie die Erforschung des humanen Gedächtnisses.

Beispielhaft sollen nachfolgend zwei Aspekte vertieft werden, die derzeit in der aktuellen Diskussion sind und die sicherlich in den nächsten Jahren in der Diagnostik an Bedeutung gewinnen werden, da sie ein großes Potenzial für das ambulante Assessment in sich bergen. Dies ist zum einen die Verwendung von RFID-Chips (Abschnitt 8.1.1) und zum anderen die Möglichkeiten einer Lokalisierung und Bewegungskontrolle (Abschnitt 8.1.2).

### 8.1.1 RFID-Chips

RFID steht für *Radio Frequency Identification*. Bei dieser Technik, die 1940 für den militärischen Bereich entwickelt wurde und seit 1977 für die zivile Nutzung freigegeben ist, handelt es sich um millimetergroße Chips, die heutzutage in Etiketten, Gegenständen oder Verpackungen eingelassen sind oder dort angebracht werden. Diese Chips sind mit einem Radiofrequenzmodul verbunden. Chip und Modul werden „Transponder“, „RFID-tag“ oder kurz „tag“ genannt. Werden elektromagnetische Wellen von einem Lesegerät eingebracht, dann sendet der Chip – durch das Prinzip der magnetischen Induktion angeregt – einen klar identifizierbaren Code an ein Lesegerät. Der Code kann aus binären Daten (z. B. bezahlt vs. nicht bezahlt) oder komplexeren Daten wie z. B. der bisherigen Bewegungshistorie bestehen.

Das Lesegerät liest die auf dem Transponder gespeicherten Daten aus, ein angeschlossenes System entschlüsselt diese und vergleicht sie mit bereits abgelegten Daten. Eine Ortung und ein „Auslesen“ dieser Chips ist auch durch die Kleidung oder Taschen hindurch möglich. In Abhängigkeit von der Antenne, der Funkfrequenz und den Umweltbedingungen können bei dieser passiven Funkversion, d. h. die Energie wird erst durch das Lesegerät erzeugt, bis zu 15 Meter überwunden werden (Bundesamt für Sicherheit in der Informationstechnik, 2004). Bei aktiven Chips, die mit Batterien ausgestattet sind (aktive Transponder) können auch mehrere Kilometer überwunden werden. Bauweise und Funktionsweise der verschiedenen Chipmodelle sind sehr heterogen. Als definierende Merkmale und kleinster gemeinsamer Nenner werden daher genannt:

(1) Elektronische Identifikation: Das System ermöglicht eine eindeutige Kennzeichnung von Objekten durch elektronisch gespeicherte Daten. (2) Kontaktlose Datenübertragung: Die Daten können zur Identifikation des Objekts drahtlos über einen Funkfrequenzkanal ausgelesen werden. (3) Senden auf Abruf (on call): Ein gekennzeichnetes Objekt sendet seine Daten nur dann, wenn ein dafür vorgesehenes Lesegerät diesen Vorgang abrufen (Bundesamt für Sicherheit in der Informationstechnik, 2004, S. 15).

Ursprünglich aus Sicherheitsgründen in Ausweisen und anderen Identitätspapieren eingebaut, wird diese Methode heute beispielsweise zur Warenverfolgung in der Logistik eingesetzt. Weiterhin werden Tags zur Ortung, Kennzeichnung und zur Messung von Bewegungsmustern von Tieren eingesetzt. Aufgrund bisheriger Pilotstudien und ausgehend von Zukunftsszenarien (z. B. Ducatel, Bogdanowicz, Scapolo, Leijten & Burgelman, 2001) sind auch für die psychologische Diagnostik die im Folgenden beschriebenen Einsatzfelder denkbar.

*Kauf- und Nutzungsverhalten von Produkten.* Durch die Einbindung von Tags in Produkten ist eine verhaltensnahe Messung des Kaufverhaltens und des Austauschs von Gütern möglich. Diese Methodik ist – den „informed consent“ der Testpersonen vorausgesetzt – in der Markt- und Werbeforschung anwendbar. Da die Chips in Verpackungen auch nach dem Gebrauch einlesbar sind, ist es auch möglich, den Verbleib von Verpackungen im Müll zu verfolgen, was beispielsweise für die Diagnostik in der Umweltpsychologie nutzbar ist. Die meist binären Daten, die im RFID-Chips gespeichert sind („Ist der Chip an Ort XY: Ja oder Nein“), beziehen sich zunächst jedoch nur auf die Bewegungen der Chips. Diese können auf unterschiedliche Verhaltensweisen und Ursachen zurückzuführen sein. Daher sind ergänzende psychometrische Erhebungsmethoden notwendig.

Eine Anwendung von „Smart-Cards“ beschreiben Lambert, Plumb und Looise (2005). Es bleibt unklar, ob es sich hierbei um den Einsatz von RFID-Chips handelt, gleiches wäre aber durchaus mit der RFID-Technologie möglich. Mithilfe dieser Smart-Card wurde das Essverhalten von Schülern in einer Schulcafeteria gemessen. Auf der Karte waren Angaben zur Personen (Alter, Geschlecht, Name) gespeichert. Jedes Menü, das in der Cafeteria auswählbar war, hatte einen speziellen Code, dem Angaben zu Kaloriengehalt, Portionsgröße und andere Nährwertanalysen zugeordnet waren. Zusätzlich wurde die Menge an zurückgegebenem Essen über die Smart-Card erfasst. Die Auswertung der Daten auf der Ebene von Gruppenmittelwerten ermöglichte es, Essenstrends zu identifizieren und anhand des Unterschieds zwischen Essenswahl und Essenskonsum auch einen Maßstab über die Qualität des Essens zu erhalten. Ebenso erlaubt diese Technologie auch eine personenbezogene Diagnostik. So konnten

der Kalorienverbrauch, Schwankungen in der Nährwertzusammensetzung, die längsschnittlichen Essenspräferenzen und Aussagen über das Essverhalten (Uhrzeit, Menge, Kalorien) ermittelt werden. Ähnliche Auswertung des Kundenverhaltens sind heute schon mit den vielfältigen Kundenbindungsprogrammen (z. B. Payback etc.) möglich.

*Direkte Datenerhebung und Datenspeicherung.* Neben einer einfachen Detektion der Chips mittels Lesegerät sind auch heute schon Chips im Einsatz, die sensorische Daten (z. B. Feuchtigkeit und Belastung von Gütern) messen. Einsatzorte in der Psychologie, bei denen mittels solcher Sensoren unauffällig physiologische Parameter gemessen werden, sind ebenso denkbar.

### 8.1.2 Tracking und Lokalisierung

Als Tracking wird die Aufzeichnung von Bewegungsdaten von Gütern, Tieren oder Menschen bezeichnet. Die Überwachung der Bewegungsmuster wird entweder durch RFID-Chips oder aber durch die Ortungen mittel GPS ermöglicht. Mit dem geplanten Einsatz des europäischen Satellitensystem GALILEO wird das bisherige „Global Positioning System“ (GPS) ergänzt und damit auch eine genauere Lokalisation möglich sein. Neben der Ortung von RFID-Chips sind heute schon Lokalisierung via Handy oder anhand von WLAN-Zugangspunkten möglich (vgl. Mattern, 2005). Ebenso ist eine Erfassung von Bewegungs- und Handlungsdaten in entsprechend präparierten Räumen für diagnostische Zwecke möglich. So können beispielsweise innerhalb eines Hauses entsprechende Sensoren an Lichtschaltern, elektrischen Geräten etc. angebracht werden, die Auskunft über das Bewegungsverhalten in den eigenen vier Wänden geben (Intille et al., 2003).

Werden RFID-Chips zur Analyse von Bewegungsdaten verwendet, dann können diese entweder in der Umgebung angebracht sein, am Körper getragen werden („wearable computing“) oder aber implantiert sein. Schon heute sind solche implantierten Chips als Zugangskontrollen für Diskotheken oder als Zugang zum eigenen häuslichen Bereich anstelle eines Schlüssels möglich („embedded computing“; vgl. <http://www.foebud.org/rfid/wo-gibt-es-rfid#>). Verschiedene Firmen (z. B. e-foxx oder Kidsfinder) bieten entsprechende Ortungstechnik an. Bekannt wurde das Tracking von Personen im Zusammenhang mit der Diskussion um die „elektronische Fußfessel“. Die Nutzung dieser Technologien zur „elektronischen Aufenthaltsüberwachung“ ist seit dem 01. Januar 2011 durch das Gesetz zur Neuordnung des Rechts der Sicherungsverwahrung auch in Deutschland flächendeckend möglich. Die Verfahren dienen der Kontrolle der Bewegungsmuster von Straffälligen bzw. der Kontrolle des Aufenthaltes in vorgeschriebenen Räumlichkeiten (Mayer, 2004). Allein in England wurden zwischen

1999 und 2005 150.000 Straffällige mit dieser Methode überwacht (Nellis, 2005). Begründet wird der Einsatz mit der Möglichkeit zur Kontrolle des Aufenthaltsortes, der besseren Möglichkeit der Resozialisierung in der gewohnten Umgebung und der Informationssammlung über das Verhalten von Straftätern. Letztlich ist mit einer solchen Kontrolle auch ein spezifisches „Profiling“ möglich: Welcher Täter hält sich wie lange wo auf?

Visionäre Vorarbeiten für diese Technologie stammen vom amerikanischen Psychiater Schwitzgebel (1964, 1967), der in den 60er Jahren von der Idee beseelt war, dass die Kriminalität durch eine umfassende Überwachung in den Griff zu bekommen sei. Die damals in Harvard entwickelte Sender-Empfänger-Kombination war mehrere Kilogramm schwer und ermöglichte die Ortung in einem Abstand von 400 Meter. Die „Schwitzgebel-Maschine“ wurde 1969 patentiert. Das Instrumentarium konnte sich zunächst nicht richtig in der Praxis durchsetzen, doch entwickelte Schwitzgebel weitere Ideen, wie z. B. die telemetrische Übermittlung von physiologischen Signalen und die ambulante Erhebung des Blutalkohol-Wertes.

Wie schnell die Grenzen zwischen Messung und Überwachung verschwinden, macht das Beispiel des „electronic tagging“ bei Dementen deutlich. Das orientierungslose Herumwandern findet sich häufig bei diesen Personen. Mittels eines Fußbandes kann der Patient lokalisiert werden. Ebenso ist es aber möglich, durch die Aufstellung mehrerer Messgeräte das Verlaufsmuster oder die Häufigkeit von Bewegungen zu erfassen. Schon der Einsatz solcher Mittel als Alternative zu mechanischen oder pharmakologischen Maßnahmen gegen den Bewegungsdrang ist eine schwierige Entscheidung (Hughes & Louw, 2002). Hinsichtlich der Nutzung solcher Daten für die Diagnostik sind folglich noch strengere Richtlinien anzusetzen, als dies bei einer rein therapeutischen oder präventiven Nutzung der Fall ist.

## 8.2 Ambulantes Assessment als Beitrag zur diagnostischen Grundlagenforschung

Die oben angesprochenen Entwicklungen werden zu einer verstärkten Diagnostik im Feld führen. Die dabei gewonnenen Datensätze sind nicht mit den Daten vergleichbar, die bei der Anwendung klassischer Fragebögen entstehen. Die Anzahl an Datenpunkten ist groß und die Varianz der Daten wird durch situative Einflüsse mitbestimmt. Messungen *in situ* verschieben den Fokus von Traits auf States. Die Messmethode erlaubt es, den Einfluss von Kontextreizen und die gegenseitige Beeinflussung unterschiedlicher Traitfacetten bei verschiedenen Personen zu analysieren.

Während es umfangreiche Literatur zur Traitdekomensation gibt, gibt es wenig Forschung zur „state decompensation“. Zur Analyse von Einflussfaktoren sind Messungen auf verschiedenen Ebenen (das sog. „concurrent assessment“) notwendig. Zur Differenzierung der Einflussfaktoren können selbst wieder technische Instrumentarien verwendet werden. So wird zur Separierung von emotionalen und motorischen Aspekten bei der Messung physiologischer Variablen beispielsweise die Akzelerometrie verwendet.

Gerade durch den Vergleich klassischer Messmethoden mit den neuen Messmethoden ergeben sich Chancen für die Diagnostik. Relevante Fragestellungen sind beispielsweise: (a) Lassen sich die Konstrukte, die bisher nur mit Fragebögen gemessen wurden, auch mit den neuen technischen Möglichkeiten bestätigen? Führen die Messungen zu gleichen Ergebnissen? (b) Ist es möglich, anhand des Vergleiches von klassischen Fragebogenergebnissen und ambulantem Assessment Subgruppen zu identifizieren, die in unterschiedlichem Maße eine Übereinstimmung zeigen? Entsprechendes versuchen Stone et al. (1999) im Hinblick auf die Schmerzmessung. (c) Welche Faktoren beeinflussen die Äquivalenz? (d) Geht die Messungsstrategie mit einer spezifischen Reaktivität einher? Denkbar sind z. B. Einflüsse durch das technische Know-How oder durch die Einstellung gegenüber modernen diagnostischen Verfahren. (e) Wie ist die Akzeptanz und Compliance generell und in verschiedenen Testgruppen?

Neben der Auswertung auf der Ebene der Items bietet eine technische Diagnostik oft einen Mehrwert gegenüber Fragebögen. Bei der Messungen mittels PDA auf ein Prompting hin fallen beispielsweise Daten zur Reaktions- und Latenzzeiten an. Durch die mehrfachen Messungen sind längsschnittliche Auswertungen, Erhebung der Stabilität oder die Analyse von Antwortstilen möglich. Diese „Zusatzdaten“ müssen jedoch in entsprechende theoretische Modelle eingeordnet werden. Wie bei der Diskussion zur Auswertungsstrategie des ambulanten Assessments deutlich wurde, stellen die Daten besondere Anforderungen an die statistische Auswertung. Schon die Messung physiologischer Daten macht deutlich, welche Schwierigkeiten aufkommen können, wenn aus unzähligen Daten sinnvolle Parameter gewonnen werden sollen. Um die Einflüsse von Personen (State und Trait), Situationen und Interaktionen dieser Aspekte zu kontrollieren, sind besondere statistische Methoden notwendig (Schwartz & Stone, 1998; West & Hepworth, 1991). Die Weiterentwicklung des „data mining“ wird zukünftig eine wichtige Aufgabe sein, die nur in Zusammenarbeit mit Informatikern zu lösen ist. Nicht nur im Hinblick auf die „Beschreibungsprobleme“ (Deskription, Abweichungsanalyse, Datenreduktion), sondern auch im Hinblick auf „Prognoseprobleme“ (prognostische Validität, Zeitreihenanalyse) bedarf es methodischer Weiterentwicklungen (Fayyad, Piatetsky-Shapiro, Smyth & Uthurusamy, 1996).

### 8.3 Automatisierte Verbindung von Diagnostik und Intervention

Durch den Einsatz technischer Geräte beim ambulanten Assessment ist es nicht nur möglich zeitnahe Daten zu diagnostischen Zwecken aufzunehmen, sondern diese auch auszuwerten und daraus unmittelbare Interventionen abzuleiten. Letztlich ist schon die Darstellung der Herzfrequenz auf einer „Pulsuhr“ nicht nur eine Datenerhebung, sondern auch das Ergebnis einer Datenauswertung, denn inmitten der vielen Störsignale, die vom Sensor empfangen werden, müssen die relevanten Daten „gefiltert“ und schließlich in einen durchschnittlichen Zahlenwert umgerechnet werden.

Wird das Ergebnis solcher Auswertungen direkt an die Patienten oder Klienten weitergegeben und auch hierzu passende therapeutische Empfehlungen abgeleitet, spricht Leopold (2001) von „Theranostic“, Heron und Smyth (2010) nennen dies „Ecological momentary interventions“. Eine Insulinpumpe ist hierfür ein bekanntes Beispiel: Die Messung des Blutzuckerwertes führt zur direkten Umsetzung einer Interventionen. Weitere Beispiele findet man in der Patientenedukation, z. B. bei der automatisierten Ableitung von Diätmaßnahmen bei Übergewicht nach vorherigem Ernährungs-Assessment (Agras, Taylor, Feldman, Losch & Burnett, 1990) oder beim Management der Asthmaerkrankung auf der Grundlage respiratorischer Variablen (Leopold & Schandry, 2001). Im Bereich der Luftfahrt ist die Diagnostik der Vigilanz mittels des Electronic Pilot-Activity Monitors (PAM) möglich. Dieses Gerät misst ausgehend von der bisherigen Flugdauer aufgrund bestimmter EEG- oder EOG-Aktivitäten (EOG: Elektrokulogramm) das Ausmaß der Vigilanz und animiert durch ein Warnsignal dazu, entsprechende Schlummerpausen einzulegen (Cabon, Mollard, Coblentz & Speyer, 2003).

Neue Möglichkeiten die Diagnostik und Therapie verbinden ergeben sich auch durch Fernsehkanäle mit Feedback-Funktion (DVB-RCS-Technik). Was zunächst nur für Bestellsysteme in Home-Shopping-Kanälen, zur Abstimmung und für Gewinnspiele gedacht war, ermöglicht auch eine Anwendung für den Gesundheitssektor. Erste Anwendung schildert der Artikel von Vouyioukas, Maglogiannis und Pasiadis (2007).

Es ist schon jetzt absehbar, dass die Stufen des diagnostischen Prozesses (Datenerfassung, Auswertung und Nutzung) in Zukunft verschmelzen und die Rolle der beteiligten Instanzen (Messgerät, Testleiter, Testperson) neu definiert werden müssen. Zurzeit findet die Auswertung bei den Testleitern oder – im Falle des self-monitoring – bei den Testpersonen zusammen. Diese entscheiden dann über die Wertigkeit der Ergebnisse und die Umsetzung möglicher Interventionen, die sich aus der Diagnostik ergeben. Das Beispiel der Insulinpumpe verdeutlicht, dass diese Aufgabenteilung weitgehend aufgehoben

ist. Hier erfolgt die Messung automatisiert, ebenso die Auswertung und auch die nachfolgende Gegenregulation des Blutzuckerwertes mittels Insulin. Diagnostik und Intervention funktionieren autonom, sind jedoch jederzeit durch die Person beeinflussbar.

Eine Weiterentwicklung dieser Idee findet man in der Konzeption des „intelligent home“ oder „smart home“. Hier übernimmt die Technik selbst die Umsetzung der Messungen in Aktionen und Entscheidungen („empowerment technology“; Intille et al., 2003). Ein Beispiel dafür sind „smarte“ Kühlschränke, die mittels RFID-Technologie erkennen, welche Produkte noch vorrätig sind und entsprechende Informationen an den Hausbewohner weitergeben („control technology“) oder aber automatisiert Waren nachbestellen („empowerment technology“). Die Bestellung von Produkten ist aufgrund von Präferenzen des Kühlschranksbesitzers möglich, die durch ambulantes Assessment in Kombination mit Variablen wie Raumtemperatur, Anzahl der Personen im Haushalt etc. ermittelt werden. „Smarte“ Gegenstände können sich merken, wo sie schon überall waren und sich kontextbezogen verhalten, z. B. kontextsensitiv weitere Datenaufzeichnungen bewirken. Eine durchdachte Kombination solcher intelligenten Instrumente, die zudem untereinander vernetzt sind, stellt dann das ambient home oder smart home dar.

*Ambient Intelligence* (AMI) wird im Technologieprogramm der „European Community Information Society“ (ISTAG) definiert als „the convergence of ubiquitous computing, ubiquitous communication, and interfaces adapting to the user“ (Raisinghani et al., 2004). Sensoren und Chips stehen dabei miteinander in Verbindung, gleichen untereinander Daten ab, nutzen zur Bewertung bisherige Datenspeicherungen um dann entsprechende Aktuatoren in Gang zu setzen.

Solche intelligenten Umgebungen sollen eine Erleichterung bei Aktivitäten des täglichen Lebens wie Kochen oder Einkaufen ermöglichen. Erste Modellprojekte finden sich im Bereich des altengerechten Wohnens, wo viele Aktivitäten der Personen durch die Routine der technischen Geräte übernommen werden. Die Arbeitsgruppe AmSys der Universität Kaiserslautern (vgl. <http://www.amsys.uni-kl.de>) entwickelt im Teilbereich „Assisted Living“ beispielsweise Möglichkeiten, um Vitaldaten von älteren Personen zu messen, die im Falle einer kritischer Überschreitung eine Rückmeldung oder Alarmierung auslösen. Ebenso sind automatisierte Abschaltungen von Elektrogeräten bei Gefahr möglich.

Grundlage für die intelligente Umgebungsgestaltung sind vorausgehende diagnostische Prozesse mittels technischer Gerätschaften, die unauffällig im Felde integriert sind. An die Stelle klassischer Eingabevorrichtungen an stationären Computern wie Mäuse, Joysticks oder Tastaturen treten nun unauffällig Messungen von Bewegungen, Gerüchen, Formen und Gestaltungen oder auch

Sprachmustern. Entsprechende Sensoren, die „intelligent social user interfaces“ (ISUIs), sind ubiquitär vorhanden („ubiquitous computing“). Weil die Informationsverarbeitung in immer mehr Bereichen unseres täglichen Lebens Einfluss erhält und viele Alltagsgegenstände einbezieht, wird von „pervasive computing“ gesprochen. Die Sammlung von Daten und deren Auswertung verläuft ebenfalls auf vielfältige Weise („ubiquitous communication“), z. B. mittels drahtloser Datenübertragung.

Zur allgegenwärtigen Kommunikation gehört auch die Vernetzung von Sensoren in *Sensornetzen*. Dabei werden beispielsweise Temperatur-, Feuchtigkeits- oder in Zukunft physiologische Daten gemessen und untereinander ausgetauscht. Für die Zukunft sind ebenfalls Übertragungen von Daten über soziale Netzwerkstrukturen denkbar, die dann genauere Informationen über die Gewohnheiten, Einstellungen und Verhaltensweisen beinhalten (Ribaud, 2003).

Diese Messungen, die bei „smart technologies“ automatisiert ausgewertet und in Aktuatoren umgesetzt werden, sind auch innerhalb der Psychologie für diagnostische Zwecke nutzbar. So bediente sich die Studie von Boyce und Geller (2001) der „Smart-car“-Technologie: Die in manchen Autos schon eingebauten „in-vehicle information systems“ (IVIS) können dem Autofahrer eine Entscheidungshilfe beim Fahren sein oder in kritischen Situationen selbstständig in die Steuerung des Fahrzeugs eingreifen. Bekannt ist die Kollisions-Vermeidung-Vorrichtung, die automatisch aktiviert wird, wenn der Abstand zum vorangehenden Fahrzeug zu gering ist, und eine Bremsung veranlasst. Ein weiteres Beispiel ist die Messung der Vigilanz des Fahrers durch die videounterstützte Kontrolle der Augenlider: Wird eine kritische Schwelle im Hinblick auf den Augenlidschluss unterschritten, dann ertönt ein Warnsignal, dass das weitere Einschlafen des Fahrers verhindert. Die diagnostischen Prozesse, die im „smart car“ direkt in Gegensteuerungen umgewandelt werden, lassen sich auch zur Diagnose des Fahrverhaltens nutzen. Durch die unauffällige Erfassung von Augen- und Kopfbewegungen, Bremsverhalten, Bewegung des Fahrzeugs in verschiedenen Ebenen, Geschwindigkeit und Messung des Abstandes zum vorher fahrenden Fahrzeug etc. gelingt es im Feld, Kennzeichen von gefährlichem Fahrverhalten zu identifizieren. Die Erfassung dieser „smart data“ hat Vorteile gegenüber der verzögerten Erfassung anhand von Verkehrsdelikten und Unfallzahlen oder der Erhebung des Fahrverhaltens mittels Fragebogen.

#### 8.4 Risiken und Probleme zukünftiger Entwicklungen

Den offensichtlichen Möglichkeiten und Vorteilen technischer Instrumentarien im ambulanten Assessment stehen aber auch erkennbare *Probleme und Gefahren* gegenüber. Die Vorstellung allgegenwärtiger unauffälliger Chips und

entsprechender Messgeräte sowie die wachsende Anzahl an diagnostischen Möglichkeiten lassen Ängste an einen Orwellschen Überwachungsstaat aufkommen. Dass dies nicht unbegründet ist, zeigen die Darstellungen von Albrecht und McIntyre (2005) in ihrem Buch „Spychips“. Die möglichen Risiken fasst der folgende Satz zusammen: „The old sayings that ‚the wall have ears‘ and ‚if these walls could talk‘ have become the disturbing reality. The world is filled with all-knowing, all reporting things“ (Lucky, 1999).

Im Gegensatz zur Computertestung, der man sich durch das Entfernen vom Testplatz entziehen kann, oder der ambulanten physiologischen Messung, die man durch Abnahme der Elektroden beenden kann, ist man beim Smart-home-Konzept mit einer Vielzahl von unauffälligen Sensoren umgeben, denen man sich schwer entziehen kann und es aufgrund der Annehmlichkeiten vielleicht auch gar nicht mehr will. Die *Minimalisierung und Einbindung ins Alltagsleben*, die „livability functions“, wie Hufford et al. (2001) sie nennen, macht deutlich, dass sich die ursprünglich reaktive Diagnostik der nicht reaktiven Diagnostik annähert. Während der klassische Fragebogen eine klar erkennbare Erhebungsmethode ist und man sich einer Testung entziehen kann, besteht beim Ambient-intelligence-Konzept, die Gefahr, dass ohne Kenntnis und – aufgrund der benutzerfreundlichen Interfaces auch unreflektiert – Daten erhoben werden. Diese Technologie hat nicht nur Auswirkungen für den häuslichen Bereich, sondern auch für die Arbeitswelt. „Organizations are exploring ways to incorporate gaze tracking, emotion detection, speech recognition, and gesture recognition in a ‚smart‘ computer, that in the future will often be invisible, with an intuitive user interface“ (Raisinghani et al., 2004).

Langheinrich, Coroama, Bohn und Mattern (2005) systematisieren die Durchdringung von einstmalig „messfreien“ Räumen anhand folgender Dimensionen: natürliche Grenzen werden überwunden, z. B. Mauern, Türen, Kleider; soziale Grenzen werden überwunden, z. B. findet eine Erhebung auch von anderen Personen über Personen statt, z. B. bei der Analyse von Arztbesuchen etc.; zeitliche und örtliche Grenzen fallen weg, d. h. Verhalten aus früheren Zeiten beeinflusst die Bewertung heutiger Daten und das Verhalten an einem Ort beeinflusst die Datenerhebung an einem anderen Ort.

Durch umfassende Erhebungs- und Speichermöglichkeiten fallen die Grenzen der beschränkten Nachhaltigkeit („ephemeral or transitory effects“) weg, d. h. kurzfristige Äußerungen, flapsige Bemerkungen oder Ähnliches, die sonst schnell vergessen sind, bleiben langfristig zugänglich und wirken nach. Die mit den neuen technischen Möglichkeiten einhergehenden Bedrohungen der Privatsphäre beziehen sich auf zwei Dimensionen der Privatheit: Zum einen auf die „data privacy“, also die Frage des Zugangs zu persönlichen Daten, und zum

anderen auf die „location privacy“. Damit ist der Zugang fremder Personen zu persönlichen Bewegungsmustern und Orten gemeint. Die Sammlung von Bewegungsmustern eines Objektes (z. B. einer Reisetasche) ist per se nicht problematisch, schwierig wird dies erst dann, wenn deutlich wird, wem die Tasche gehört. Durch die Speicherung von Daten auf Objekten („Dinge mit Gedächtnis“) und die gleichzeitige Erhebung von personenbezogenen Daten im Hinblick auf diese Dinge ist also die Privatheit bedroht.

Erhebungen zu den Einstellungen von RFID-Chips zeigen, dass das Bewusstsein für die Bedrohung der Privatheit durch Möglichkeiten der ubiquitären und unaufdringlichen Diagnostik eher gering ausgeprägt ist. Nur eine kleine Gruppe der „extreme rejecters“ setzt sich vehement gegen die Nutzung ein, während in der breiten Bevölkerung kaum eine kritische Auseinandersetzung mit den Verfahren stattfindet (Rothensee & Spiekermann, 2008).

Bedrohungen der Privatsphäre können sich dabei nicht nur durch das Ausspähen der Daten, sondern auch durch die Einspeisung falscher Daten ergeben, die dann für die Person mit möglichen Nachteilen verbunden ist. Da innerhalb der Psychologie die Technik zukünftig vermehrt genutzt wird, ist es wichtig, die Sorgen der Bevölkerung und speziell der Testpersonen ernst zunehmen und die *Balance zwischen relevanten Erhebungsinteressen und dem Schutz der Privatsphäre* durch eine reflektierte Anwendung zu erreichen.

Insbesondere in der Anwendung der Methode im klinischen Bereich ist es wichtig, dass Patienten den Anwendern und der Technik selbst trauen können. Vertrauensbildende Maßnahmen sind beispielsweise: (1) der faire und transparente Gebrauch der Daten, (2) die Möglichkeit Zugang zu den Daten zu erhalten und den Zugang durch Dritte zu beschränken, sowie (3) die Anonymität der Daten zu sichern, wenn es sich um Gruppentestungen handelt (Sillence & Briggs, 2008). Mehr noch als bei den ambulanten physiologischen Messmethoden, bei denen wenigstens die Anwesenheit der Elektroden an die Messung erinnert, ermöglichen die RFID-Chips eine Datenerhebung ohne Sicht- und Berührungskontakt. Die Anwendung geht daher mit strengen Anforderungen an die forschungsethischen Bestimmungen und die entsprechende Instruktion der Testpersonen einher (vgl. <http://www.rfid-handbook.de>; Bundesamt für Sicherheit in der Informationstechnik, 2004). Wenn auf der einen Seite umfassende Daten erfasst werden und auf der anderen Seite Messabsichten und die Verwendung der Daten für die Testpersonen unklar bleiben, dann handelt es sich um eine *Informationsasymmetrie*, die es durch umfassende Aufklärung der Testpersonen zu beheben gilt.

In Zukunft werden intelligente Systeme entstehen, die beispielsweise eine nutzerabhängige Preisfestlegung („pay-per-use“) ermöglichen. So wird der

Rechnungsbetrag für die Autoversicherung aufgrund der Dauer und der Art der Autonutzung berechnet werden. Die Diagnose der Nutzungshäufigkeit und -art wird dann auf die Rechnungsstellung Einfluss haben und die pauschale Bemessung ersetzen. Grundlage für diese adaptiven Bemessungen werden ebenfalls diagnostische Prozesse sein, z. B. „smart technologies“, die das Fahrverhalten erfassen und aufzeichnen. Für einige entsteht dadurch eine Kostenersparnis, für andere eine finanzielle Mehrbelastung. Was auf den ersten Blick wie eine Maßnahme wirkt, die zu mehr Gerechtigkeit führt, stellt sich in letzter Konsequenz als Aufgabe des Solidaritätsprinzips dar, das keine sozialen Nischen für Problemfälle ermöglicht.

In Anbetracht der rasanten Entwicklungen und der unglaublichen Möglichkeiten dieser Technologie ist es auch aus ethischen Gründen wichtig, für die „De-Diagnostizierung“ bestimmter Lebensbereiche einzutreten – auch wenn es der Psychologie schwer fällt. Die Jagd nach diagnostischer Sicherheit und inkrementeller Varianzaufklärung durch neue diagnostische Möglichkeiten muss dort ihre Grenzen finden, wo kein Rückzug der Testpersonen mehr möglich ist und das Prinzip „Freiwilligkeit“ für große Bevölkerungsschichten aufgehoben wird. Die Bevölkerung ist für Verletzung der Privatsphäre zwar sensibel, allerdings auch schnell beeinflussbar, wenn finanzielle Vorteile winken. Die Einführung von *Kundenkarten* (z. B. Payback-Karten) hat gezeigt, dass viele Personen – trotz anfänglicher Bedenken – diese Karten nutzen. Mehr als 50 % der Bevölkerung haben wenigstens *eine* Kundenkarte (Stand 2002; Bundesamt für Sicherheit in der Informationstechnik, 2004). Dies zeigt, dass finanzielle Anreize mögliche Probleme schnell vergessen lassen.

Die *rechtlichen Bestimmungen* für die Messung im Felde sind nach Meinung von Experten im Bundesdatenschutzgesetz ausreichend geregelt. Dort ist festgehalten, dass die Speicherung von Daten der Zustimmung der betroffenen Person bedarf. Ebenso ist auf die Verwendung von RFID-Chips und möglicher Datenspeicherung hinzuweisen und entsprechend das Einverständnis einzuholen. Im Falle der RFID-Technologie müssen Möglichkeiten geschaffen werden, aus der „Überwachung“ auszusteigen. Hierzu sind sogenannte „clipped tags“ empfehlenswert, die es den Verbrauchern bzw. Testpersonen erlauben, manuell die Antenne des Clips abzubrechen und damit die Datenerhebung zu beenden. Eine umfassende Darstellung des Datenschutzes bei der Anwendung von Tags findet sich bei Langheinrich et al. (2005).

Experten sehen im Hinblick auf das „ubiquitous computing“ noch einen Regelungsbedarf, da erst die Kombination verschiedener Daten aus unterschiedlichen Quellen die Privatsphäre bedroht. Besonders für die Übertragung sensibler Daten auf drahtlosem Wege sind die Sicherheitstechniken (Verschlüsselungen) weiterzuentwickeln.

## 9 Schlussfolgerungen

Die Ausführungen haben deutlich gemacht, dass die Entwicklung technischer Geräte neue Möglichkeiten für die Messungen in der natürlichen Umgebung ermöglicht. Eine verhaltensnahe Messung im Felde ist zur Aufklärung der relevanten psychologischen Konstrukte von hohem Wert und wird schon deshalb in den nächsten Jahren an Bedeutung gewinnen. Folgt man den Worten von Baumeister et al. (2007), dann ist es dringend erforderlich, dass sich die Psychologie wieder auf ihre ursprünglichen Motive besinnt und sich um die Aufklärung menschlichen Verhaltens bemüht. Fragebögen haben dort ihre Berechtigung, wo es darum geht Einstellungen oder subjektive Repräsentationen von Erleben und Verhalten zu erfassen. Solche Daten sind aber nicht zwingend verhaltensvalid und sollten daher durch unmittelbare Messmethoden ergänzt werden. Besonders dann, wenn die Erfassung von Art, Häufigkeit, Kontext oder Dauer des Verhaltens für gezielte Interventionen wichtig ist, ist das ambulante Assessment eine bessere Strategie. Mithilfe der technischen Entwicklungen, insbesondere der Minimalisierung der Messgeräte und den verbesserten Möglichkeiten der Datentransfers ergeben sich nun neue Chancen, das menschliche Verhalten in der natürlichen Umgebung wieder zum Gegenstand der Forschung zu machen. Der Einsatz von PDAs und Handys zur Diagnostik kann inzwischen als umfangreich beforschtes Verfahren bewertet werden, dem in der Ausbildung und Anwendung aber eine stärkere Rolle zukommen sollte.

Für zukünftige Entwicklungen wie beispielweise den Einsatz von RFID-Chips ist es wichtig, dass die Psychologie aktiv gestaltend bei der Entwicklung für diagnostische Zwecke mitwirkt (Raento, Qulasvirta & Eagle, 2009). Zwei Gründe sollen dazu genannt werden: (1) Für viele technische Geräte mangelt es heute noch an diagnostischen Verfahren, um die Einstellung, die Bedienung und mögliche Nebenwirkungen zu erfassen. Hier muss die Psychologie vermehrt geeignete Testinstrumente entwickeln. (2) Für den Einsatz technischer Mittel in der Diagnostik ist es wichtig, dass die Inhalte die Wahl der technischen Geräte mitbestimmt. Es darf nicht darum gehen, ein trendiges Diagnostikum zu schaffen, sondern Technik zu nutzen, um auch messtheoretische Vorschritte zu machen oder aber um neue psychologische Konstrukte zu erschließen, die mit bisherigen Verfahren nicht (geeignet) erfasst werden konnten.

## Literatur

- Agras, W. S., Taylor, C. B., Feldman, D. E., Losch, M. & Burnett, K. (1990). Developing computer-assisted therapy for the treatment of obesity. *Behavior Therapy*, 21, 99–109.
- Albrecht, K. & McIntyre, L. (2005). *Spychips: How major corporations and government plan to track your every move with RFID*. Nashville, TN: Nelson Current.

- Alpers, G. W., Wilhelm, F. H. & Roth, W. T. (2005). Psychophysiological assessment during exposure in driving phobic patients. *Journal of Abnormal Psychology*, 114 (1), 126–139.
- Aserinsky, E. & Kleitman, N. (1953). Regularly occurring periods of eye mobility and concomitant phenomena during sleep. *Science*, 118, 273–274.
- Bachofen, M., Nakagawa, A., Marks, I. M., Park, J. M., Greist, J. H. & Baer, L. (1999). Home self-assessment and self-treatment of obsessive-compulsive disorder using a manual and a computer-conducted telephone interview: Replication of a UK-US study. *Journal of Clinical Psychiatry*, 60, 545–549.
- Bailey, S., Martin, C. S., Lynch, K. G. & Pollock, N. K. (2000). Reliability and concurrent validity of DSM-IV subclinical symptom ratings for alcohol use disorders among adolescents. *Alcoholism: Clinical and Experimental Research*, 24, 1795–1802.
- Barker, E. T., Williams, R. L. & Galambos, N. L. (2006). Daily spillover to and from binge eating in first-year university females. *Eating Disorders*, 14 (3), 229–242.
- Barker, R. G. (1968). *Ecological psychology. Concepts and methods for studying the environment of human behavior*. Stanford, CA: Stanford University Press.
- Barker, R. G. & Wright, H. (1951). *One boy's day. A specimen record of behavior*. New York: Harper & Brothers.
- Barker, R. G. & Wright, H. (1955). *Midwest and its children. The psychological ecology of an American town*. New York: Harper & Brothers.
- Barrett, D. J. & Feldman-Barrett, L. (2000). *The Experience-Sampling Program (ESP)*. Boston, MA: Boston College.
- Baumeister, R. F., Vohs, K. D. & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, 2 (4), 396–403.
- Beal, D. J. & Weiss, H. M. (2003). Methods of ecological momentary assessment in organizational research. *Organizational Research Methods*, 6 (4), 440–464.
- Bennett, G. G., Wolin, K. Y., Puleo, E. & Emmons, K. M. (2006). Pedometer-determined physical activity among multiethnic low-income housing residents. *Medicine & Science in Sports & Exercise*, 38 (4), 768–773.
- Ben-Zeev, D., Young M. A. & Madsen J. W. (2009). Retrospective recall of affect in clinically depressed individuals and controls. *Cognition and Emotion*, 23 (5), 1021–1040.
- Berger, A. M., Wielgus, K. K., Young-McCaughan, S., Fischer, P., Farr, L. & Lee, K. A. (2008). Methodological challenges when using actigraphy in research. *Journal of Pain and Symptom Management*, 36 (2), 191–199.
- Bersoff, D. M. & Bersoff, D. N. (2000). Ethical issues in the collection of self-report data. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman & V. S. Cain (Eds.), *The science of self-report. Implications for research and practice* (pp. 9–24). Mahwah, NJ: Erlbaum.
- Bolger, N., Davis, A. & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, 54, 579–616.
- Boseck, J. J., Engel, S. G., Allison, K. C., Crosby, R. D., Mitchell, J. E. & De Zwaan, M. (2007). The application of Ecological Momentary Assessment to the study of the night eating. *International Journal of Eating Disorders*, 40 (3), 271–276.

- Bouten, C. V., Westertep, K. R. & Verduin, M. (1994). Assessment of energy expenditure for physical activity using triaxial accelerometer. *Medical Science in Sports and Exercises*, 26, 1516–1523.
- Boyce, T. E. & Geller, E. S. (2001). A technology to measure multiple driving behaviors without self-report of participant reactivity. *Journal of Applied Behavior Analysis*, 34 (1), 39–55.
- Brewin, C. R., Andrews, B. & Gotlib, I. H. (1993). Psychopathology and early experience. A reappraisal of retrospective reports. *Psychological Bulletin*, 111 (1), 82–98.
- Buck, R. & Morley, S. (2006). A daily process design study of attentional pain control strategies in the self-management of cancer pain. *European Journal of Pain*, 10 (5), 385–398.
- Bundesamt für Sicherheit in der Informationstechnik (2004). *Risiken und Chancen des Einsatzes von RFID-Systemen*. Bonn: Bundesamt für Sicherheit in der Informationstechnik.
- Burke, L. E. (2001). Electronic measurement. In L. E. Burke & I. Ockene (Eds.), *Compliance in healthcare and research* (pp. 117–138). Armonk, NY: Futura Publishing.
- Buse, L. & Pawlik, K. (1996). Ambulatory behavioral assessment and in-field performance testing. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 29–50). Seattle, WA: Hogrefe & Huber.
- Buse, L. & Pawlik, K. (2001). Computer-assisted ambulatory performance tests in everyday situations: Construction, evaluation, and psychometric properties of a test battery measuring mental activation. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 3–24). Seattle, WA: Hogrefe & Huber.
- Bussmann, H., van den Berg-Emons, R. & Stam, H. (2001). Posture and motion analysis by ambulatory accelerometry. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 193–206). Seattle, WA: Hogrefe & Huber.
- Bussmann, J. B., Martens, W. L. J., Tulen, J. H. M., Schasfoort, F. C., van den Berg-Emons, H. J. G. & Stam, H. J. (2001). Measuring daily behavior using ambulatory accelerometry: The Activity Monitor. *Behavior Research Methods, Instruments & Computers*, 33, 349–356.
- Cabon, P., Mollard, R., Coblenz, A. & Speyer, J.-J. (2003). Electronic pilot-activity monitor: A countermeasure against fatigue on long-haul flights. *Aviation, Space, and Environmental Medicine*, 74, 679–682.
- Cattell, R. B. (1946). *The description and measurement of personality*. New York: World Book.
- Cerin, E. & Barnett, A. (2006). A processual analysis of basic emotions and sources of concerns as they are lived before and after a competition. *Psychology of Sport and Exercise*, 7 (3), 287–307.
- Chandra, S., Shiffman, S. & Scharf, D. M. (2007). Daily smoking patterns: Their determinants, and implications for quitting. *Experimental and Clinical Psychopharmacology*, 15 (1), 67–80.
- Collins, R. L., Kashdan, T. B. & Gollnisch, G. (2003). The feasibility of using cellular phones to collect ecological momentary assessment data: Application to alcohol consumption. *Experimental and Clinical Psychopharmacology*, 11 (1), 73–78.
- Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, 57, 505–528.
- Corkey, R. & Parkinson, L. (2002). Interactive voice response: Review of studies, 1989–2000. *Behavior Research Methods, Instruments & Computers*, 34, 342–353.
- de Vries, S. I., Bakker, I., Hopman-Rock, M., Hirasing, R. A. & van Mechelen, W. (2006). Clinimetric review of motion sensors in children and adolescents. *Journal of Clinical Epidemiology*, 59 (7), 670–680.
- Delichatsios, H. K., Freidman, R. H., Glanz, K., Tennstedt, S., Smigelski, C. & Pinto, B. M. (2001). Randomized trial of a „talking computer“ to improve adults eating habits. *American Journal of Health Promotion*, 15, 215–224.
- DeLongis, A., Hemphill, K. J. & Lehmann, D. R. (1992). A structured diary methodology for the study of daily events. In F. Bryant, J. Edwards, R. S. Tinsdale, E. J. Posavac, L. Heath, E. Henderson & Y. Suarez-Balcazar (Eds.), *Methodological issues in applied social psychology. Social psychological applications to social issues* (Vol. 2, pp. 83–109). New York: Plenum.
- DeVries, M. W. (Ed.). (1992). *The experience of psychopathology. Investigating mental disorders in their natural setting*. Cambridge: Cambridge University Press.
- Dimsdale, J. E., Ancoli-Israel, S., Ayalon, L., Elsmore, T. F. & Gruen, W. (2007). Taking fatigue seriously, II: variability in fatigue levels in cancer patients. *Psychosomatics*, 48 (3), 247–252.
- Ducatel, L., Bogdanowicz, M., Scapolo, F., Leijten, J. & Burgelman, J. C. (2001, October). *That's what friends are for. Ambient intelligence (AmI) and the IS in 2010*. Paper presented at the Congress of Innovations for an e-Society, Challenges for Technology Assessment Berlin, 17–19 October.
- Dunbar-Jacob, J., Sereika, S., Rohay, J. & Burke, L. E. (1998). Electronic methods in assessing adherence to medical regimens. In D. S. Krantz & A. Baum (Eds.), *Technology and methods in behavioral medicine* (pp. 95–113). Mahwah, NJ: Erlbaum.
- Ebner-Priemer, U. W. & Kubiak, T. (2007). Psychological and psychophysiological ambulatory monitoring. *European Journal of Psychological Assessment*, 23 (4), 214–226.
- Ebner-Priemer, U. W., Kuo, J. & Welch, S. S. (2006). A valence-dependent group-specific recall bias of retrospective self-reports: A study of Borderline personality disorder in everyday life. *Journal of Nervous and Mental Disease*, 194 (10), 774–779.
- Eguchi, K., Pickering, T. G., Hoshida, S., Ishikawa, J., Ishikawa, S., Schwartz, J. E. et al. (2008). Ambulatory blood pressure is a better marker than clinic blood pressure in predicting cardiovascular events in patients with/without type 2 diabetes. *American Journal of Hypertension*, 21, 443–450.
- Erdfelder, E. & Funke, J. (Hrsg.). (2004). *Allgemeine Psychologie und deduktivistische Methodologie*. Göttingen: Vandenhoeck & Ruprecht.
- Fahrenberg, J. (1996). Ambulatory assessment: Issues and perspectives. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 3–20). Seattle, WA: Hogrefe & Huber.
- Fahrenberg, J. (2001). Origins and developments of ambulatory monitoring and assessment. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 587–614). Seattle, WA: Hogrefe & Huber.
- Fahrenberg, J., Foerster, F. & Müller, W. (1996). Laboratory and field studies for improvement of ambulatory monitoring methodology. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory*

- assessment. *Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 257–268). Seattle, WA: Hogrefe & Huber.
- Fahrenberg, J. & Myrtek, M. (1996). *Ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies*. Ashland, OH: Hogrefe & Huber.
- Fahrenberg, J. & Myrtek, M. (Eds.). (2001). *Progress in ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies*. Ashland, OH: Hogrefe & Huber.
- Fahrenberg, J., Myrtek, M., Pawlik, K. & Perrez, M. (2007). Ambulantes Assessment – Verhalten im Alltagskontext erfassen. *Psychologische Rundschau*, 58 (1), 12–23.
- Farmner, A., Gibson, O. J., Tarassenko, L. & Neil, A. (2005). A systematic review of telemedicine interventions to support blood glucose self-monitoring in diabetes. *Diabetic Medicine*, 22, 1372–1378.
- Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P. & Uthurusamy, R. (1996). *Advances in knowledge discovery and data mining*. Cambridge, MA: MIT Press.
- Ferrante, F. (2005). Evolving telemedicine/eHealth technology. *Telemedicine Journal and e-Health*, 11, 370–383.
- Field, M. J. (1997). Telemedicine: A guide to assessing telecommunications in healthcare. *Journal of Digital Imaging*, 10, 28.
- Fitzgibbon, A. & Reiter, E. (2003). „Memories for life“ *Managing information over a human lifetime*. Grand Challenge proposal. Aberdeen: UK Computing Research Committee (UKCRC).
- Flechter, L. A., Erickson, D., J., Toomey, T. L. & Wagenaar, A. C. (2003). Handheld computers: A feasible alternative to paper forms for field data collection. *Evaluation Review*, 27, 165–178.
- Fleeson, W. (2001). Toward a structure-and-process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, 73, 1107–1117.
- Fox, K. R., Stathi, A., McKenna, J. & Davis, M. G. (2007). Physical activity and mental well-being in older people participating in the Better Ageing Project. *European Journal of Applied Physiology*, 100 (5), 591–602.
- Freedman, M. J., Lester, K. M., McNamara, C., Milby, J. B. & Schumacher, J. E. (2006). Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. *Journal of Substance Abuse Treatment*, 30 (2), 105–111.
- Friedberg, F. & Sohl, S. J. (2009). Longitudinal change in chronic fatigue syndrome: What home-based assessments reveal? *Journal of Behavioral Medicine*, 32, 209–221.
- Funder, D. C. (2001). Personality. *Annual Review of Psychology*, 52, 197–221.
- Gallagher, J. M., Stewart, T. V., Pathak, P. K., Werner, J. J., Dickinson, L. M. & Hickner, J. M. (2008). Data collection outcomes comparing paper forms with PDA forms in an office-based patient survey. *Annals of Family Medicine*, 6, 154–160.
- Gamberini, L., Cardullo, S., Seraglia, B. & Bordin, A. (2009). Neuropsychological testing through a Nintendo Wii. *Annual Review of CyberTherapy and Telemedicine*, 8, Special issue: Imaging the future, 22–25.
- Gee, P., Coventry, K. R. & Birkenhead, D. (2005). Mood state and gambling: Using mobile telephones to track emotions. *British Journal of Psychology*, 96 (1), 53–66.

- Gemmell, J., Bell, G. & Lueder, R. (2006). MyLifeBits: A personal database for everything. *Communications of the ACM*, 49, 88–95.
- Gloster, A. T., Richard, D. C., Himle, J., Koch, E., Anson, H., Lokers, L. et al. (2008). Accuracy of retrospective memory and covariation estimation in patients with obsessive-compulsive disorder. *Behavior Research and Therapy*, 46 (5), 642–655.
- Godaert, G. L. R., Sorbi, M. J., Peters, M., Dekkers, C. & Geenen, R. (2001). Ambulatory monitoring of diurnal changes in pain in chronic pain disorder, migraine, and rheumatoid arthritis. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 123–128). Seattle, WA: Hogrefe & Huber.
- Goldberg, J., Wolf, A., Silberstein, S., Gebeline-Myers, C., Hopkins, M., Einhorn, K. et al. (2007). Evaluation of an electronic diary as a diagnostic tool to study headache and premenstrual symptoms in migraineurs. *Headache*, 47 (3), 384–396.
- Goulet, G., Mongrain, V., Desrosiers, C., Paquet, J. & Dumont, M. (2007). Daily light exposure in morning-type and evening-type individuals. *Journal of Biological Rhythms*, 22 (2), 151–158.
- Hacker, E. D. & Ferrans, C. E. (2007). Ecological momentary assessment of fatigue in patients receiving intensive cancer therapy. *Journal of Pain and Symptom Management*, 33 (3), 267–275.
- Hank, P. & Schwenkmezger, P. (1996). Computer-assisted versus paper-and-pencil-based self-monitoring: An analysis of experimental and psychometric equivalence. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 85–100). Seattle, WA: Hogrefe & Huber.
- Hank, P., Schwenkmezger, P. & Schumann, J. (2001). Daily mood reports in hindsight: Results of a computer-assisted time sampling study. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 143–156). Seattle, WA: Hogrefe & Huber.
- Hanson, E. K. S. & Barzanski, J. (2001). Ecological momentary assessment in aviation: The development of a pilot experience evaluating device (PEED) for the in-flight registration of flight phases, mental effort, and reaction time. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 477–492). Seattle, WA: Hogrefe & Huber.
- Hedeker, D., Mermelstein, R. J. & Demirtas, H. (2008). An application of a mixed-effects location scale model for analysis of Ecological Momentary Assessment (EMA) data. *Biometrics*, 64 (2), 627–634.
- Hedges, S. M., Krantz, D. S., Contrada, R. J. & Rozanski, A. R. (1990). Development of a diary for use with ambulatory monitoring of mood, activities, and physiological function. *Journal of Psychopathology and Behavioral Assessment*, 12, 203–217.
- Heger, R. (1990). Apparative Datenerfassung in Feldstudien. *Diagnostica*, 36, 60–80.
- Heron, K. E. & Smyth, J. M. (2010). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology*, 15 (Pt 1):1–39. Epub 2009 Jul 28.
- Heilman, K. J. & Porges, S. W. (2007). Accuracy of the LifeShirt(R) (Vivometrics) in the detection of cardiac rhythms. *Biological Psychology*, 75 (3), 300–305.

- Ho, J. & Intille, S.S. (2005). *Using context-aware computing to reduce the perceived burden of interruptions from mobile devices*. Paper presented at the CHI 2005 – Conference on Human Factors in Computing Systems, New York.
- Hopper, J.W., Su, Z. & Looby, A.R. (2006). Incidence and patterns of polydrug use and craving for ecstasy in regular ecstasy users: An ecological momentary assessment study. *Drug and Alcohol Dependence*, 85 (3), 221–235.
- Hoppmann, C.A. & Riediger, M. (2009). Ambulatory assessment in lifespan psychology: An overview of current status and new trends. *European Psychologist*, 14, 98–108.
- Houtveen, J.H. & de Geus, E.J.C. (2009). Noninvasive psychophysiological ambulatory recordings: study design and data analysis strategies. *European Psychologist*, 14, 132–141.
- Hufford, M.R., Shiffman, S., Paty, J. & Stone, A.A. (2001). Electronic momentary assessment: Real-world, real-time measurement of patient experience. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 69–92). Seattle, WA: Hogrefe & Huber.
- Hughes, J.C. & Louw, S.J. (2002). Electronic tagging of people with dementia who wander: Ethical considerations are possibly more important than practical benefits. *British Medical Journal*, 325 (7369), 847–848.
- Intille, S.S., Rondoni, J., Kukla, C., Iaconi, I. & Bao, L. (2003). *A context-aware experience sampling tool*. Paper presented at the CHI '03 Extended Abstracts on Human Factors in Computing Systems.
- Isomursu, M., Tähti, M. & Väinämö, S. (2007). Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. *International Journal of Human-Computer Studies*, 65 (4), 404–418.
- Jahng, S., Wood, P.K. & Trull, T.J. (2008). Analysis of affective instability in ecological momentary assessment: Indices using successive difference and group comparison via multilevel modeling. *Psychological Methods*, 13 (4), 354–375.
- Jain, A., Martens, W.L.J., Mutz, G., Weiß, R.K. & Stephan, E. (1996). Towards a comprehensive technology for recording and analysis of multiple physiological parameters within their behavioral and environmental context. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 237–256). Seattle, WA: Hogrefe & Huber.
- Jamison, R.N., Fanciullo, G.J. & Baird, J.C. (2004). Computerized dynamic assessment of pain: Comparison of chronic pain patients and healthy controls. *Pain Medicine*, 5 (2), 167–177.
- Jamison, R.N., Scrocco, T. & Parris, W.C.V. (1989). The influence of physical and psychosocial factors on accuracy of memory for pain in chronic pain patients. *Pain*, 37, 289–294.
- Jamner, L.D., Shapiro, D. & Alberts, J. (1998). Mood, blood pressure, and heart rate: Strategies for developing a more effective ambulatory mood diary. In D.S. Krantz & A. Baum (Eds.), *Technology and methods in behavioral medicine* (pp. 195–220). Mahwah, NJ: Erlbaum.
- Jones, J.W. & Dages, K.D. (2003). Technology trends in staffing and assessment: A practice note. *International Journal of Selection and Assessment*, 11, 247–252.
- Käppler, C., Brügger, G. & Fahrenberg, J. (2001). Pocketcomputer-unterstütztes Assessment mit MONITOR. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 22, 249–266.

- Käppler, C. & Rieder, S. (2001). Does the retrospection effect hold as a stable phenomenon? First results from a transcultural self-monitoring study of mood and cognitive states in Brazil and Germany. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 113–122). Seattle, WA: Hogrefe & Huber.
- Kazdin, A.E. (1974). Reactive self-monitoring: The effect of response desirability, goal setting and feedback. *Journal of Consulting and Clinical Psychology*, 42, 704–716.
- Kennedy, R.S. & Bittner, A. (1980). Development of performance evaluation tests for environmental research (PETER): Complex counting test. *Aviation, Space, and Environmental Medicine*, 51, 142–144.
- Kessler, D.K., Chakko, S. & Kessler, K.M. (1994). The post-holter interview. *Journal of Ambulatory Monitoring*, 7, 29–31.
- Kikuchi, H., Yoshiuchi, K., Miyasaka, N., Ohashi, K., Yamamoto, Y., Kumano, H. et al. (2006). Reliability of recalled self-report on headache intensity: investigation using ecological momentary assessment technique. *Cephalalgia*, 26 (11), 1335–1343.
- Kimhy, D., Delespaul, P. & Corcoran, C. (2006). Computerized experience sampling method (ESMc): Assessing feasibility and validity among individuals with schizophrenia. *Journal of Psychiatric Research*, 40 (3), 221–230.
- Klumb, P.L., Elfering, A. & Herre, C. (2009). Ambulatory Assessment in industrial/organizational psychology – fruitful examples and methodological issues. *European Psychologist*, 14, 120–131.
- Krantz, D.S. & Baum, A. (Eds.). (1998). *Technology and methods in behavioral medicine*. Mahwah, NJ: Erlbaum.
- Kubiak, T. & Jonas, C. (2007). Applying circular statistics to the analysis of monitoring data. *European Journal of Psychological Assessment*, 23 (4), 227–237.
- Lambert, N., Plumb, J. & Looise, B. (2005). Using smart card technology to monitor the eating habits of children in a school cafeteria: 1. Developing and validating the methodology. *Journal of Human Nutrition and Dietetics*, 18, 243–254.
- Langheinrich, M., Coroama, V., Bohn, J. & Mattern, F. (2005). Living in a smart environment – Implications for the coming ubiquitous information society. *Telecommunications Review*, 15 (1), 132–143.
- Larson, R. & Csikszentmihalyi, M. (1983). The Experience Sampling Method. *New Directions for Methodology of Social & Behavioral Science*, 15, 41–56.
- Le, B. & Hat, N. (2006). Pocket-sized psychology studies: Exploring daily diary software for Palm Pilots. *Behavior Research Methods*, 38, 325–332.
- Leopold, C. (2001). From ambulatory assessment to telemedicine: Conceptual, technical and political aspects of remote data assessment and transfer. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 569–586). Seattle, WA: Hogrefe & Huber.
- Leopold, C. & Schandry, R. (2001). Giving feedback to asthma patients. Ambulatory monitoring in patient education. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 505–524). Seattle, WA: Hogrefe & Huber.

- Lichtenberg, U., Eitmann, J. & Goldmann, G. (2003). *Behavior Settings revisited. Forschungsbericht*. Berlin: Abteilung Psychologie im Institut für Sozialwissenschaft.
- Lin, C.-T., Chen, Y.-C., Huang, T. Y., Chiu, T.-T., Ko, L. W., Liang, S. F. et al. (2008). Development of wireless brain computer interface with embedded multitask scheduling and its application on real-time driver's drowsiness detection and warning. *IEEE Transactions on Biomedical Engineering*, 55 (5), 1582–1591.
- Lucky, R. (1999). Reflections March 1999. Available from <http://www.argreenhouse.com/papers/r lucky/spectrum/connect.shtml> (05.07.2010)
- Mahlberg, R. & Walther, S. (2007). Actigraphy in agitated patients with dementia: Monitoring treatment outcomes. *Zeitschrift für Gerontologie und Geriatrie*, 40 (3), 178–184.
- Malinowski, B. (1935). *Coral gardens and their magic: Soil tilling and agricultural rites. Vol 1*. Bloomington, IN: Indiana University Press.
- Matos, S., Birring, S.S., Pavord, I.D. & Evans, D.H. (2007). An automated system for 24-h monitoring of cough frequency: the leicester cough monitor. *IEEE Transactions on Biomedical Engineering*, 54 (8), 1472–1479.
- Mattern, F. (2005). Allgegenwärtige und verschwindende Computer. *Praxis der Informationsverarbeitung und Kommunikation (PIK)*, 28 (1), 29–36.
- Mattocks, C., Leary, S., Ness, A., Deere, K., Saunders, J., Kirkby, J. et al. (2007). Intraindividual variation of objectively measured physical activity in children. *Medicine & Science in Sports & Exercise*, 39 (4), 622–629.
- Mayer, M. (2004). *Modellprojekt Elektronische Fußfessel. Wissenschaftliche Befunde zur Modellphase des hessischen Projekts. forschung aktuell – research in brief* (No. 23). Freiburg: Max-Planck-Institut für ausländisches und internationales Strafrecht.
- McDaniel, A.M., Benson, P.L., Roesener, G.H. & Martindale, J. (2005). An integrated computer-based system to support nicotine dependence treatment in primary care. *Nicotine & Tobacco Research*, 7, S57-S66.
- McKay, J.R., Franklin, T.R., Patapis, N. & Lynch, K.G. (2006). Conceptual, methodological, and analytical issues in the study of relapse. *Clinical Psychology Review*, 26 (2), 109–127.
- McMurray, R.G., Ward, D.S., Elder, J.P., Lytle, L.A., Strikmiller, P.K., Baggett, C.D. et al. (2008). Do overweight girls overreport physical activity? *American Journal of Health Behavior*, 32 (5), 538–546.
- Meegan, S.P. & Goedereis, E.A. (2006). Life task appraisals, spouse involvement in strategies, and daily affect among short- and long-term married couples. *Journal of Family Psychology*, 20 (2), 319–327.
- Mehl, M.R., Gosling, S.D. & Pennebaker, J.W. (2006). Personality in its natural habitat: manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, 90 (5), 862–877.
- Mehl, M.R. & Holleran, S.E. (2007). An empirical analysis of the obtrusiveness of and participants compliance with the electronically activated recorder (EAR). *European Journal of Psychological Assessment*, 23 (4), 248–257.
- Mehl, M.R., Pennebaker, J.W., Crow, D.M., Dabbs, J. & Price, J.H. (2001). The electronically activated recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, and Computers*, 33, 517–523.

- Merrilees, C.E., Goeke-Morey, M. & Cummings, M. (2008). Do event-contingent diaries about marital conflict change marital interactions? *Behavior Research and Therapy*, 46, 253–262.
- Middeke, M., Baumgart, P., Gotzen, R., Krönig, B., Rascher, W., Schrader, J. et al. (1992). *Ambulante Blutdruck-Langzeitmessung (ABDM)*. Stuttgart: Thieme.
- Morren, M., Dulmen, S.V., Ouwerkerk, J. & Bensing, J. (2008). Compliance with momentary pain measurement using electronic diaries: A systematic review. *European Journal of Pain*, 13, 354–365.
- Muraven, M., Collins, R.L. & Shiffman, S. (2005). Daily fluctuations in self-control demands and alcohol intake. *Psychology of Addictive Behaviors*, 19 (2), 140–147.
- Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P. & van Os, J. (2009). Experience sampling research in psychopathology: opening the black box of daily life. *Psychological Medicine*, 39 (9), 1533–1547.
- Myrtek, M. (2004). *Heart and emotion: Ambulatory monitoring studies in everyday life*. Ashland, OH: Hogrefe & Huber.
- Myrtek, M., Aschenbrenner, E. & Brügger, G. (2005). Emotions in everyday life: An ambulatory monitoring study with female students. *Biological Psychology*, 68, 237–255.
- Nagels, G., Engelborghs, S. & Vloeberghs, E. (2007). Correlation between actigraphy and nurses' observation of activity in dementia. *International Journal of Geriatric Psychiatry*, 22 (1), 84–86.
- Nellis, M. (2005). Out of this world: The advent of the satellite tracking of offenders in England and Wales. *Howard Journal of Criminal Justice*, 44 (2), 125–150.
- Nezlek, J.B. (2008). An Introduction to multilevel modeling for social and personality psychology. *Social and Personality Psychology Compass*, 8, 842–860.
- Palermo, T.M., Valenzuela, D. & Stork, P.P. (2004). A randomized trial of electronic versus paper pain diaries in children: Impact on compliance, accuracy, and acceptability. *Pain*, 107, 213–219.
- Palmblad, M. & Tiplady, B. (2004). Electronic diaries and questionnaires: Designing user interfaces that are easy for all patients to use. *Quality of Life Research*, 13, 1199–1207.
- Pambianco, G., Wing, R. & Robertson, R. (1990). Accuracy and reliability of the Caltrac accelerometer for estimating energy expenditure. *Medical Science in Sports and Exercise*, 22, 858–862.
- Pawlik, K. (1996). Contributions to differential psychology from ambulatory assessment: A researcher's plaidoyer. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 21–25). Seattle, WA: Hogrefe & Huber.
- Pawlik, K. & Buse, L. (1982). Rechnergestützte Verhaltensregistrierung im Feld. Beschreibung und erste psychometrische Überprüfung einer neuen Erhebungsmethode. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 3, 101–118.
- Perrez, M. & Reicherts, M. (1996). A computer-assisted self-monitoring procedure for assessing stress-related behavior under real life conditions. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 51–68). Seattle, WA: Hogrefe & Huber.

- Perrez, M., Wilhelm, P., Schoebi, D. & Horner, M. (2001). Simultaneous computer-assisted assessment of causal attribution and social coping in families. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 25–44). Seattle, WA: Hogrefe & Huber.
- Peters, M. L. & Crombez, G. (2007). Assessment of attention to pain using handheld computer diaries. *Pain Medicine*, 8 (Suppl. 3), 110–120.
- Pickering, T., Harshfield, G., Devereaux, R. & Laragh, J. (1985). What is the role of ambulatory blood pressure monitoring in the management of hypertensive patients? *Hypertension*, 7, 171–177.
- Pinheiro, J. C. & Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer.
- Pohl, R. F. (2004). *Cognitive illusions: A handbook on fallacies and biases in thinking, judgment and memory*. New York: Psychology Press.
- Psychology Software Tools, Inc. (1998). *AGARD Standardized Tests for Research with Environmental Stressors (STRES) Battery*. Pittsburgh, PA: Psychology Software Tools, Inc.
- Raento, M., Qulasvirta, A. & Eagle, N. (2009). Smartphones: An emerging tool for social scientists. *Sociological Methods & Research*, 37 (3), 426–454.
- Raisinghani, M. S., Benoit, A., Ding, J., Gomez, M., Gupta, K., Gusila, V. et al. (2004). Ambient intelligence: Changing forms of human-computer interaction and their social implications *Journal of Digital Information*, 5 (4), 271.
- Rau, R. (2001). Objective characteristics of jobs affect blood pressure at work, after work, and at night. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 361–386). Seattle, WA: Hogrefe & Huber.
- Rau, R. & Richter, P. (1996). Psychophysiological analysis of strain in real life work situations. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 271–286). Seattle, WA: Hogrefe & Huber.
- Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods*, 2, 173–185.
- Raudenbush, S. W., Brennan, R. T. & Barnett, R. C. (1995). A multivariate hierarchical model for studying psychological change within married couples. *Journal of Family Psychology*, 9, 161–174.
- Reisch, T., Ebner-Priemer, U. W., Tschacher, W., Bohus, M. & Linehan, M., M. (2008). Sequences of emotions in patients with borderline personality disorder. *Acta Psychiatrica Scandinavica*, 118 (1), 42–48.
- Reuschenbach, B. (2006). Ecological momentary assessment (EMA) – Anwendungen und Perspektiven in der Pflegewissenschaft. *PrInternet – angewandte Pflegeforschung*, 8 (9), 437–449.
- Reuschenbach, B. & Mahler, C. (in Druck). *Handbuch pflegebezogener Assessment-Instrumente*. Bern: Huber.
- Ribaldo, B. (2003). When the walls have ears. *Computer Technology Review*, 23 (5), 31.

- Richard, D. C. S. & Bobicz, K. (2003). Computers and behavioral assessment: Six years later. *Behavior Therapist*, 26, 219–223.
- Riis, J., Loewenstein, G., Baron, J., Jepson, C., Fagerlin, A., Ubel, P. A. et al. (2005). Ignorance of hedonic adaptation to hemodialysis: A study using ecological momentary assessment. *Journal of Experimental Psychology: General*, 134, 3–9.
- Ritterband, L. M., Cox, D. J., Kovatchev, B. P., Borowitz, S. M., Ling, W. D. & Marshall, B. (2001). A telecommunication monitoring system for clinical and research practice. *Professional Psychology: Research and Practice*, 32, 636–641.
- Roenkae, A., Malinen, K., Kinnunen, U., Tolvanen, A. & Laetsae, T. (2010). Capturing daily family dynamics via text messages: Development of the mobile diary. *Community, Work & Family*, 13 (1), 5–21.
- Roschelle, J., Penuell, W. & Yarnall, L. (2005). Handheld tools that 'Informate' assessment of student learning in science: A requirements analysis. *Journal of Computer Assisted Learning*, 21 (3), 190–203.
- Rothensee, M. & Spiekermann, S. (2008). Between extreme rejection and cautious acceptance: Consumers' reactions to RFID-based in retail. *Social Science Computer Review*, 26 (1), 75–86.
- Rowan, P. J., Cofta-Woerpel, L., Mazas, C. A., Vidrine, J. I., Reitzel, L. R., Cinciripini, P. M. et al. (2007). Evaluating reactivity to ecological momentary assessment during smoking cessation. *Experimental & Clinical Psychopharmacology*, 15 (4), 382–389.
- Rutledge, T., Stucky, E., Dollarhide, A., Shively, M., Jain, S. et al. (2009). A real-time assessment of work stress in physicians and nurses. *Health Psychology*, 28, 194–200.
- Salovey, P., Sieber, W. J., Jobe, J. B. & Willis, G. B. (1993). The recall of physical pain. In N. Schwarz & S. Sudman (Eds.), *Autobiographical memory and the validity of retrospective reports* (pp. 89–106). New York: Springer.
- Sbarra, D. A. & Ferrer, E. (2006). The structure and process of emotional experience following nonmarital relationship dissolution: Dynamic factor analyses of love, anger, and sadness. *Emotion*, 6 (2), 224–238.
- Schandry, R. & Leopold, C. (1996). Ambulatory assessment of self-monitored subjective and objective symptoms of diabetic patients. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 393–402). Seattle, WA: Hogrefe & Huber.
- Schlütz, D. & Scherer, H. (2001). Der Einsatz der 'Experience Sampling Method' in der Medienwissenschaft. *Zeitschrift für Medienpsychologie*, 13 (3), 146–149.
- Schomer, D. L. (2006). Ambulatory EEG telemetry: how good is it? *Journal of Clinical Neurophysiology*, 23 (4), 294–305.
- Schuler, H. (1982). Ethische Probleme. In J.-L. Patry (Hrsg.), *Feldforschung. Methoden und Probleme sozialwissenschaftlicher Forschung unter natürlichen Bedingungen* (S. 341–364). Bern: Huber.
- Schwarb, T. M. (1996). *Die wissenschaftliche Konstruktion der Personalauswahl*. München: Hampp.
- Schwartz, J. & Stone, A. A. (1998). Data analysis for EMA studies. *Health Psychology*, 17, 6–16.

- Schwartz, N. & Sudman, S. (1994). *Autobiographical memory and the validity of retrospective reports*. New York: Springer.
- Schwitzgebel, R. (1964). A program for research in behavior electronics. *Behavioral Science, 9*, 233–238.
- Schwitzgebel, R. (1967). Electronic innovation in the behavioral sciences: A call to responsibility. *American Psychologist, 22*, 364–370.
- Searles, J. S., Helzer, J. E. & Walter, D. E. (2000). Comparison of drinking patterns measured by daily reports and timeline follow back. *Psychology of Addictive Behavior, 14*, 277–286.
- Seibt, R., Berndt, A., Knöpfel, D. & Scheuch, K. (2001). Ambulatory 24-hour monitoring of brachial and finger blood pressure in normotensive and hypertensive males. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 293–312). Seattle, WA: Hogrefe & Huber.
- Shiffman, S. (2000). Real-time self-report of momentary states in the natural environment: Computerized ecological momentary assessment. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe & H. S. Kurtzman (Eds.), *The science of self-report: Implications for research and practice* (pp. 277–296). Mahwah, NJ: Erlbaum.
- Shiffman, S., Kirchner, T. R., Ferguson, S. G. & Scharf, D. M. (2009). Patterns of intermittent smoking: An analysis using Ecological Momentary Assessment. *Addictive Behaviors, 34*, 514–519.
- Shiffman, S., Scharf, D. M. & Shadel, W. G. (2006). Analyzing milestones in smoking cessation: Illustration in a nicotine patch trial in adult smokers. *Journal of Consulting and Clinical Psychology, 74* (2), 276–285.
- Shiffman, S. & Stone, A. A. (1998). Ecological momentary assessment: A new tool for behavioral medicine research. In D. S. Krantz & A. Baum (Eds.), *Technology and methods in behavioral medicine* (pp. 117–131). Mahwah, NJ: Erlbaum.
- Shrier, L. A., Shih, M., Hacker, L. & de Moor, C. (2007). A momentary sampling study of the affective experience following coital events in adolescents. *Journal of Adolescent Health, 40* (4), e1–e8.
- Sidiropoulou, E. (1997). Computerdiagnostik. In H. J. Fisseni (Hrsg.), *Lehrbuch der psychologischen Diagnostik* (S. 383–410). Göttingen: Hogrefe.
- Silbermann, A., Henkel, A., Müller, A. & de Zwaan, M. (2008). The application of ecological momentary assessment to the study of compulsive buying. *Psychotherapie, Psychosomatik und Medizinische Psychologie, 58* (12), 454–461.
- Sillence, E. & Briggs, P. (2008). Ubiquitous Computing – Trust Issues for a „Health“ Society. *Social Science Computer Review, 26*, 6–12.
- Simpson, T. L., Kivlahan, D. R., Bush, K. R. & McFall, M. E. (2005). Telephone self-monitoring among alcohol use disorder patients in early recovery: A randomized study of feasibility and measurement reactivity. *Drug and Alcohol Dependence, 79*, 241–250.
- Singer, J. D. & Willett, J. B. (2003). *Applied Longitudinal Data Analysis*. New York: Oxford University Press.
- Smith, W. & Safer, M. (1993). Effects of present pain level on recall of chronic pain and medication use. *Pain, 55*, 355–361.

- Sorbi, M. J., Honkoop, P. C. & Godaert, G. L. R. (1996). A signal-contingent computer diary for the assessment of psychological precedents of the migraine attack. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 403–414). Seattle, WA: Hogrefe & Huber.
- Stemmler, G. (1996). Strategies and designs in ambulatory assessment. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 258–268). Seattle, WA: Hogrefe & Huber.
- Stephan, E., Mutz, G., Feist, A. & Weiß, R. K. (2001). Some new developments in ambulatory assessment devices. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 561–568). Seattle, WA: Hogrefe & Huber.
- Stephoe, A. (2001). Ambulatory monitoring of blood pressure in daily life: A tool for investigating psychosocial processes. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 257–270). Seattle, WA: Hogrefe & Huber.
- Stinson, J. N., Petroz, G. C., Tait, G., Feldman, B. M., Streiner, D., McGrath, P. J. et al. (2006). e-Ouch: Usability testing of an electronic chronic pain diary for adolescents with arthritis. *Clinical Journal of Pain, 22* (3), 295–305.
- Stone, A. A. & Broderick, J. E. (2007). Real-Time data collection for pain: Appraisal and current status. *Pain Medicine, 8* (Suppl. 3), S85–S93.
- Stone, A. A. & Shiffman, S. (1994). Ecological Momentary Assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine, 16*, 199–202.
- Stone, A. A. & Shiffman, S. (2002). Capturing momentary, self-report data: A proposal for reporting guidelines. *Annals of Behavioral Medicine, 24*, 236–243.
- Stone, A. A., Shiffman, S. & DeVries, M. W. (1999). Ecological momentary assessment. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 26–39). New York: Russell Sage Foundation.
- Stone, A. A., Turkkan, J. S., Bachrach, C. A., Jobe, J. B., Kurtzman, H. S. & Cain, V. S. (Eds.). (2000). *The science of self-report. Implications for research and practice*. Mahwah, NJ: Erlbaum.
- Suls, J. & Martin, R. E. (1993). Daily recording and ambulatory monitoring methodologies in behavioral medicine. *Annals of Behavioral Medicine, 15*, 3–7.
- Thieden, E. (2008). Sun exposure behaviour among subgroups of the Danish population. Based on personal electronic UVR dosimetry and corresponding exposure diaries. *Danish Medical Bulletin, 55*, 47–68.
- Trull, T. J. & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. *Psychological Assessment, 21* (4), 457–462.
- Tseng, H. M., Tiplady, B., McLeod, H. & Wright, P. (1998). Computer anxiety: A comparison of pen-based personal digital assistants, conventional computer, and paper assessment of mood and performance. *British Journal of Psychology, 89*, 599–610.
- Tulen, J. H. M., Stronks, D. L., Bussmann, J. B. J., Pepplinkhuizen, L. & Passchier, J. (2000). Towards an objective quantitative assessment of daily functioning in migraine: A feasibility study. *Pain, 86*, 139–149.

- Tulen, J. H. M., Vokers, A., Stronks, D., Cavelaars, M. & Groenevela, W. (2001). Accelerometry in clinical psychophysiology. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 207–232). Seattle, WA: Hogrefe & Huber.
- Turner, C. F., Villarroel, M. A., Rogers, S. M., Eggleston, E., Ganapathi, L., Roman, A. M. et al. (2005). Reducing bias in telephone survey estimates of the prevalence of drug use: A randomized trial of telephone audio-CASI. *Addiction, 100*, 1432–1444.
- Van den Berg-Emons, R. J., Schasfoort, F. C., De Vos, L. A., Bussmann, J. B. & Stam, H. J. (2006). Impact of chronic pain on everyday physical activity. *European Journal of Pain, 11* (5), 587–593.
- Van't Hooft, M., Díaz, S. & Swan, K. (2004). Examining the potential of handheld computers: Findings from the Ohio pep project. *Journal of Educational Computing Research, 30*, 295–311.
- Vogt, J. & Kastner, M. (2001). Psychophysiological monitoring of air traffic controllers: Exploration, simulation, validation. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 455–476). Seattle, WA: Hogrefe & Huber.
- Vouyioukas, D., Maglogiannis, I. & Pasiadis, V. (2007). Pervasive E-health services using the DVB-RCS communication technology. *Journal of Medical Systems, 31* (4), 237–246.
- Waters, A. J. & Li, Y. (2008). Evaluating the utility of administering a reaction time task in an ecological momentary assessment study. *Psychopharmacology, 97*, 25–35.
- Weiss, H. M., Beal, D. J., Barros, E. & MacDermid, S. M. (2003). *Conceptualizing performance processes: A model to guide research linking quality of life and performance* (Technical report prepared for the Office of Military Community and Family Policy, Department of Defense). West Lafayette, IN: Purdue University, Department of Psychological Sciences.
- West, S. & Hepworth, J. (1991). Statistical issues in the study of temporal data: Daily experiences. *Journal of Personality and Social Psychology, 59*, 609–662.
- Whalen, C. K., Jamner, L. D., Henker, B. & Delfino, R. J. (2001). Smoking and moods in adolescents with depressive and aggressive dispositions: Evidence from survey and electronic diaries. *Health Psychology, 20*, 99–111.
- Wilhelm, F. H., Roth, W. T. & Sackner, M. A. (2003). The LifeShirt: An advanced system for ambulatory measurement of respiratory and cardiac function. *Behavior Modification, 27* (5), 671–691.
- Wilson, G. (2001). In-flight psychophysiological monitoring. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 435–454). Seattle, WA: Hogrefe & Huber.
- Wittmann, W. W. (1985). *Evaluationsforschung. Aufgaben, Probleme und Anwendungen*. Berlin: Springer.
- Zeier, H., Häseli, A. & Fischer, J. (2001). Heart rate monitoring in an academic test situation. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 387–398). Seattle, WA: Hogrefe & Huber.

## 13. Kapitel

Einsatz technischer Mittel  
in der psychologischen Diagnostik

Joachim Funke und Bernd Reuschenbach

## 1 Einleitung

Der Einsatz technischer Mittel innerhalb der Diagnostik ist so alt wie die wissenschaftliche Disziplin selbst. Bekannt dürften apparative Verfahren zur Berufseignung sein, die zu Zeiten der „Industriellen Psychotechnik“ (Giese, 1927; Moede, 1928) Anfang des 20. Jahrhunderts eingesetzt wurden (vgl. Albert & Gundlach, 1997). Schulte (1921) entwickelte etwa eine Apparatur zur Tauglichkeitsprüfung angehender Damenfriseurinnen, mit der die Zugstärke beim Kämmen von Perücken als Indikator für die Sanftheit der Behandlung gemessen werden konnte. Über verschiedene, historisch interessante psychotechnische Prüfverfahren für Straßenbahner und Eisenbahner berichtet im Detail der Artikel von Stapf (1997).

Zur Eingrenzung des Themas ist es wichtig, zunächst den Begriff der Technik zu reflektieren. Technik (altgriechisch: *τεχνη*) bedeutet übersetzt „Fähigkeit“ oder „Kunstherrlichkeit“. Diagnostik (altgriechisch: *διαγνωσις*) bedeutet übersetzt „Entscheidung“. So gesehen kann auch die Zuweisung von Körperkonstitutionen zu bestimmten Persönlichkeitseigenschaften oder die Entwicklung von Fragebögen als Kunstherrlichkeit und damit im weitesten Sinn als „technische Diagnostik“ verstanden werden.

Wenn von *technischen Mitteln* die Rede ist, dann deutet dies auf eine weitere Facette des Technikbegriffs hin, nämlich auf die Nutzung von Mitteln, die mechanischer oder elektrotechnischer Natur sind bzw. sich entsprechendes technologisches Wissen zunutze machen. Die verwendeten Mittel sind das Ergebnis von Entwicklungsarbeiten im Bereich der Elektrotechnik, des Ma-

# ENZYKLOPÄDIE DER PSYCHOLOGIE

In Verbindung mit der  
Deutschen Gesellschaft für Psychologie

herausgegeben von

Prof. Dr. Niels Birbaumer, Tübingen  
Prof. Dr. Dieter Frey, München  
Prof. Dr. Julius Kuhl, Osnabrück  
Prof. Dr. Wolfgang Schneider, Würzburg  
Prof. Dr. Ralf Schwarzer, Berlin

Themenbereich B

Methodologie und Methoden

Serie II

Psychologische Diagnostik

Band 3

Leistungs-, Intelligenz-  
und Verhaltensdiagnostik



Hogrefe • Verlag für Psychologie  
Göttingen • Bern • Toronto • Seattle

# Leistungs-, Intelligenz- und Verhaltensdiagnostik

herausgegeben von

Prof. em. Dr. Lutz F. Hornke, Aachen  
Prof. em. Dr. Manfred Amelang, Heidelberg  
Prof. Dr. Martin Kersting, Münster



Hogrefe • Verlag für Psychologie  
Göttingen • Bern • Toronto • Seattle