

# **7 years of experience using Personal Digital Assistants (PDA) for psychometric diagnostics in 6000 inpatients and polyclinic patients**

Rose M, Walter OB, Fliege H, Becker J, Hess V, Klapp BF  
rose@charite.de  
Germany

**Abstract:** In clinical psychosomatics there is a need for adequate documentation of psychometric data both within the framework of individual diagnostics and in care-specific quality assurance. The data collection method should meet current psychometric standards, while at the same time allowing a fast and economical evaluation. In contrast to conventional methods using paper and pencil, computer-assisted testing offers considerable advantages in both regards.

Using data collected from 6434 patients between the years of 1995 and 2002, we investigated what effects the complete conversion of a psychometric basis assessment to a mobile, computer-assisted method had in terms of acceptance, data structure, and economy. We implemented the following questionnaires, among others: the Center for Epidemiologic Studies-Depression (CES-D) Scale, the Giessen Complaint Inventory (Giessener Beschwerdebogen GBB), the Berlin Mood Questionnaire (Berliner Stimmungsfragebogen BSF), the State-Trait-Anxiety Inventory (STAI), the Short-Form 36 (SF-36), the WHOQOL-Bref, the Giessen Test (GTS), the Narcissism Inventory (NI), and parts of the Psychosomatic Basis Documentation (PsyBaDo).

Based on our experience, we can conclude that

1. testing by means of mobile PDA's enjoys a high rate of acceptance among patients,
2. the organization of data is considerably improved, which guarantees an immediate evaluation for clinical and scientific lines of questioning,
3. no significant alterations result in terms of the psychometric characteristics of the questionnaire, and
4. mobile computerized testing is two-thirds cheaper than conventional testing using paper and pencil.

**Keywords:** PDA, psychometric diagnostics, computer assisted testing, economic advantages, test equivalence

## **1 Introduction**

In clinical psychosomatics and other fields aiming to improve the subjective wellbeing of patients, there is an increasing need for the adequate documentation of psychometric data, both within the framework of individual diagnostics and for general quality assurance.

As a rule, psychometric measurement is currently performed using standardized instruments in paper and pencil form. For the most part, however, this type of testing brings with it the problem that the evaluation of the instruments often takes place weeks after the actual testing is done. As a consequence, the results of the measurements are not available to the attending physician or therapist at the time when they would be most useful for the therapeutic process. Moreover, only in rare cases is the patient, who often spends hours giving information about himself, informed about the test results, nor does he receive an opportunity to discuss the results with his physician or therapist. Consequently the psychological diagnostics fall way behind the somatic diagnostics in terms of logistics, when we consider that the results of, say, blood tests are as a rule available the following day. In our opinion, the delay between testing and evaluation, above all, explains why psychometric diagnostics have previously experienced such a low level of acceptance by patients and attending physicians and therapists alike, as well as explaining the previously low level of implementation.

From our experience, the time between testing and evaluation results from the fact that entering data by hand and archiving questionnaires is very labor-intensive and therefore quite costly. Over the past several years there have been attempts from the commercial side at least to avoid the entry of data by hand. Companies such as SPSS-Munich have sold programs like their Teleform-Software since the early 90's that makes it possible to scan questionnaires. To do this, automated scanning equipment must be available, on the one hand, and the forms must be very carefully filled out, on the other. According to our experience, the latter is in practice hardly the case, thereby rendering the savings in labor and maintenance time minimal. In the attempt to principally avoid the process of retyping data and its additional danger of reading errors, commercial computer assisted systems have been on offer for several years now. These systems make it possible to fill out questionnaires directly on a PC in an economical fashion. They then can immediately evaluate the data entered, additionally guaranteeing optimized data administration (for example the Hogrefe Testsystem). However, the previous systems run, as a rule, on a desktop or laptop PC and are designed such that testing essentially takes place under the direct guidance of a trained personnel. Moreover, there is hardly any leeway for scientifically motivated modifications to the instruments.

Inspired by the methods of psychophysiological field research [Fah94], we decided in 1994 to develop our own software solution which makes mobile collection of psychometric data as well as an immediate individual and scientific evaluation of that data possible.

## **2 Software solution**

Since January 1995, we have used a software solution in the Psychosomatic Department of the Charité at the Humboldt University Berlin which allows to administer 15 different questionnaires on Psion3 PDA's and now on Revo PDA's (approximately  $16.5 \times 8.8 \times 2.3$ cm, 280g) in a flexible manner. In 1997 the largest psychosomatic clinic in Berlin, the Theodor-Wenzel-Werk (director: Dr. Keller), adopted our system for their routine diagnostics. Moreover, our system is being used within the framework of a multi-center project

carried out in cooperation with the Robert Koch Institute.

## **2.1 Logistic design**

Before the patient is examined, a staff member (secretary, intern, or data operator) first enters the patient's identification data (patient number, name, sex) into the mobile computer and selects the desired questionnaire or a pre-installed routine set of various questionnaires. The prepared PDA is handed over to the patient afterwards. The questions are self-explanatory. Responses are made by the patient on an independent basis and at any location. The PDA's used can store the responses of over 50 patients. Once the PDA is returned to the department, the data is transferred to the department's server and filed in an MS-Access database.

## **2.2 Individual application**

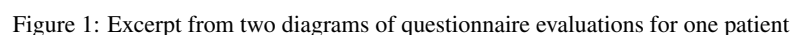
Immediately after data input by the patient, the examination data are compiled in form of a diagram by a staff assistant for the entire course of treatment or for a certain examination time point. In this way, the results of the psychometric diagnostics are, as a rule, available even before the initial contact with the attending therapist, and they can be used for the interview as needed (Fig. 1). The diagram is either created in MS-Access or in MS-Excel, which allows the user to modify the evaluation functions of the system independently.

## **2.3 Scientific application**

The transferred data are stored in a relational database in MS-Access on the department's server. In this way the data is available to all authorized users on the campus, up to date at all times, and available at each individual work place. The data organization allows for a simple link to other treatment-relevant data stored in the same database, such as diagnosis, treatment time point, duration of treatment, physician or therapist, as well as sociodemographic or somatic data. Since MS-Access can easily communicate with commonly used statistics programs, the evaluation can be carried out, according to the preference of the user, using SAS, SPSS, Statistica etc., by means of ODBC or after exporting the data.

## **3 Acceptance by the patients**

During the past 7 years, we were able to experience the work with 6434 patients and up to 15 different questionnaires within the framework of our polyclinic, concilliary, and inpatient diagnostics as well as through various scientific studies. Altogether, over 80,000



questionnaires were administered using PDA's.

From the introduction of the diagnostic method in 1995, the acceptance rate of the examination technology was generally very high among the patients. The university context might have aided this, yet we had the impression that patients generally associated the testing by the "computer" with a feeling of being held in high regard or that the diagnostics were scientifically founded, in contrast to the tests on paper that patients encounter in pulp magazines.

An initial problem was the limited use of PDA's in Germany, which meant that the curious patients did want to try out other functions of the PDA's during the examination. This, however, was dealt with optimizing the configuration of the instrument and our software, so nowadays PDA's cannot be used in any other way. With an increase in the wide-spread public availability of PDA's and an increase in the experience of department staff with operating them, difficulties concerning the technique vanished. Only the use by visually impaired patients is limited due to the small size of the display ( $12.7 \times 4.5$  cm). Less than 1% of the patients reject the examination, which corresponds to our experiences with the paper version [RFH00, SRD00, FRA01].

Our experiences correspond to those of other work groups, namely that computerized testing usually meets with a high level of acceptance even among mentally ill patients [Git90, FB89, HG86].

## **4 Acceptance by physicians and therapists**

### **4.1 Individual application**

The question which role psychometric diagnostics should take principally in therapy needs to be answered in conjunction with the question, what technical aids are to be used. In this discussion the different psychotherapeutic school plays an important role. In the treatment team at our clinic, comprising psychoanalysts, individual psychologists, behavioral psychologists, and creative therapists, do view psychometry for diagnostics quite differently. However, the measurement of patient's subjective condition for purposes of assessing the success of medical as well as psychotherapeutic treatment is met with general acceptance. This great approval can largely be traced back to the distinct reduction in workload for staff and above all to the immediate availability, transparency of data, and the high rate of acceptance by the patients. In this respect we may say that, according to our experiences, the computer assisted testing tends to have an advantage over the paper and pencil version, too.

## 4.2 Scientific application

In the utilization of data for scientific or other grouped statistical data, computer assisted testing demonstrates a number of decisive advantages:

1. the absence of missing data,
2. the avoidance of data transfer errors,
3. the uniform organization in a well known database, which affords colleagues with limited computer skills easy access and use of familiar statistics programs,
4. joint collection of psychometric and other patient-related data, and
5. access from every workplace to the current database of each day.

On the other hand, there is the disadvantage that in multi-center studies, for instance, the appropriate hardware solution has to be installed in all locations. For small numbers of tests, the paper application certainly remains the better alternative.

## 5 Psychometric characteristics

The meaning of presentation form for value is, as far as previous studies suggest, apparently dependent on the psychological construct under investigation. Thus, worse results are usually found, for instance, in intelligence tests (progressive matrices) administered on a computer compared with paper versions, leading Kubinger and Farkas [KF91] to the assumption that "the computer versions lead to hasty responses", while written work is carried out with greater care - and therefore contains fewer mistakes.

Until now measurement of personality traits appears to be largely independent of the presentation form ([Fow85, Hon88, Jäg90, SH93], overview in [Fra97]). Franke [Fra97] was able to prove the equality of paper-pencil and computer versions in student samples for the Freiburger Personality Inventory (FPI), and Bader et al. [BHK93] in an experimental design have proven the transferability of the norm values for the Giessen Test Self-Image Version. Additionally, in one of our investigations we found no differences in mean value between the data collection forms of the GTS [RHS99].

In contrast, the situation appears to present itself quite differently in more state-related instruments. Schwenkmezger and Hank [SH93] point out that far fewer investigations have been carried out in this area. Yet, the existing ones nevertheless allow the conclusion that the structure of data apparently remains intact in terms of distribution, stability, etc. but that the scale values depend on the form of data collection; in part, to an extent that holds some meaning for individual diagnostics.

As an example of this instance, we mention the characteristics of the Berlin Mood Questionnaire (BSF, [HK93]) in the paper-pencil version (gathered between 1989 and 1994) as compared with the computer assisted version implemented between 1995 and 2001.

Here we find a nearly identical factor structure (Tab.1) and internal consistency of the scales. We already demonstrated in part of the sample that one of the six scales differed between both sample groups [RHS99]. This effect was significant due to the large size of the samples, yet only made up 13% of the standard deviation and therefore is hardly likely to be clinically relevant. The observed effect already loses its significance level following an a-Bonferoni adjustment. Nevertheless, in our opinion one should follow the recommendation put forth in 1986 by national and international psychological associations (Test Curatorium of the Federation of German Psychologist Associations 1986, [Ame86, Fra97, SH93]) that one cannot assume the equivalence of different presentation forms, above all in situation-specific measurement instruments, without prior review. In any event, it seems useful to rely on, wherever possible, the most current comparative group matched to patient in terms of age and sex and measured using the same presentation form for individual evaluation of psychometric data (see [Kub95]).

## **6 Economic aspects**

Lienert & Raatz [LR94] define economy as a secondary criteria for the quality of a test regarding low consumption of material, simplicity of operation, as well as fast and easy analysis, among others. As early as 30 years ago, Kleinmuntz and McLean [KM68] pointed out the advantages of computer testing with regard to saving of time, objectivity, and flexibility. To assess the relevance of these advantages we attempted to weigh up the saved labor costs against the costs of the implemented hardware. In 1999 our calculations showed decisive savings in costs through the introduction of PDA assisted diagnostics [RHS99]. If we take into account investment costs and corresponding depreciation, at least 1.15 EUR would have to be spent per questionnaire in the paper and pencil version as opposed to 0.37 EUR using PDA testing. With continued increases in personnel costs and reductions in hardware costs, and availability of the computers for other applications, the cost of the PDA application is likely to have decreased further.

## **7 Perspectives**

In our experience, the collection of psychometric data by means of PDA's is distinctly superior to data collection by means of paper and pencil, both in terms of acceptance, ease of evaluation, and economy. For these advantages to come to fruition, the following aspects has to be guaranteed:

1. independence of location,
2. independence of staff's working hours from patient's speed in filling in forms,
3. an immediate individual evaluation of instruments, and
4. a flexible adaptation of the system even by less-experienced users.

factor variance explanation %	Paper-pencil tests n=674							PDA assisted tests n=6434						
	I 30.3	II 7.5	III 7.4	IV 13.1	V 3.4	VI 4.5	$h^2$	I 36.0	II 10.9	III 6.9	IV 6.1	V 3.2	VI 4.5	$h^2$
exhausted	<b>.84</b>	.24	.16	-.10	-.03	.10	.81	<b>.82</b>	.27	.21	-.14	-.06	.13	.83
languid	<b>.83</b>	.25	.17	-.11	-.06	.04	.79	<b>.80</b>	.21	.21	-.14	-.08	.10	.76
listless	<b>.81</b>	.20	.22	-.10	-.10	.09	.77	<b>.82</b>	.21	.23	-.14	-.07	.13	.81
tired	<b>.80</b>	.13	.14	-.08	-.03	.11	.69	<b>.80</b>	.13	.20	-.09	-.09	.16	.73
weary	<b>.77</b>	.21	.17	-.10	.06	.08	.68	<b>.77</b>	.28	.14	-.11	-.04	.13	.71
worried	.30	<b>.72</b>	.09	-.16	-.01	.24	.69	.26	<b>.76</b>	.15	-.18	.03	.20	.74
anxious	.25	<b>.61</b>	.06	-.14	.19	.25	.55	.19	<b>.73</b>	.09	-.16	.11	.12	.63
insecure	.19	<b>.69</b>	.25	-.11	-.18	.13	.62	.19	<b>.63</b>	.36	-.18	-.13	.11	.62
edgy	.20	<b>.60</b>	-.02	-.05	-.02	.33	.51	.20	<b>.62</b>	.09	.01	-.06	.27	.51
distressed	.34	<b>.62</b>	.26	-.33	-.02	.17	.70	.39	<b>.57</b>	.34	-.29	-.05	.16	.70
depressed	.35	<b>.61</b>	.31	-.30	-.02	.20	.71	.32	<b>.57</b>	.40	-.28	-.10	.22	.72
apathetic	.28	.15	<b>.73</b>	-.11	-.18	.09	.68	.26	.23	<b>.77</b>	-.08	-.14	.10	.74
uninvolved	.18	.08	<b>.76</b>	-.09	-.12	.06	.64	.22	.17	<b>.75</b>	-.05	-.11	.15	.67
indifferent	.20	.01	<b>.74</b>	.01	-.12	.06	.60	.17	.02	<b>.72</b>	-.08	-.10	.14	.58
uninterested	.17	.11	<b>.76</b>	-.05	-.20	.11	.66	.20	.20	<b>.70</b>	.00	-.17	.11	.60
bored	.00	.12	<b>.60</b>	-.09	.07	.18	.41	.05	.10	<b>.65</b>	-.05	.01	.22	.49
cheerful	.10	-.06	-.13	<b>.85</b>	.16	.06	.78	-.13	-.13	-.06	<b>.85</b>	.23	.00	.82
happy	.10	-.12	-.13	<b>.86</b>	.14	.04	.80	-.13	-.13	-.07	<b>.84</b>	.23	.01	.80
merry	-.07	-.11	-.14	<b>.87</b>	.14	.09	.81	-.15	-.18	-.10	<b>.80</b>	.20	-.03	.81
balanced	-.14	-.26	.02	<b>.58</b>	.31	-.13	.53	-.18	-.41	-.08	<b>.50</b>	.40	-.16	.63
relaxed	-.18	-.31	.08	<b>.61</b>	.30	-.06	.60	-.22	-.40	.00	.46	.39	-.09	.57
understanding	.01	.28	-.02	.44	<b>.52</b>	-.23	.59	.06	.03	-.07	.23	<b>.71</b>	-.14	.58
attentive	-.06	.02	-.21	.23	<b>.74</b>	.01	.64	-.18	-.10	-.26	.18	<b>.71</b>	-.06	.65
deliberate	.09	-.03	-.10	.14	<b>.73</b>	.14	.58	-.04	.00	.02	.04	<b>.67</b>	.13	.46
concentrated	-.16	-.21	-.18	.20	<b>.70</b>	.03	.63	-.31	-.26	-.18	.13	<b>.60</b>	-.07	.58
helpful	-.01	.40	-.03	.34	<b>.50</b>	-.20	.55	.05	.14	-.10	.24	<b>.64</b>	-.05	.50
belligerent	.04	.10	.05	.10	.05	<b>.82</b>	.70	.11	.07	.14	.02	-.03	<b>.85</b>	.75
aggressive	.06	.14	.13	-.01	.00	<b>.84</b>	.73	.13	.16	.23	-.03	-.07	<b>.83</b>	.79
mad	.13	.27	.20	-.02	.01	<b>.72</b>	.65	.14	.29	.20	-.03	.01	<b>.75</b>	.70
irritable	.18	.34	.18	-.07	-.05	<b>.71</b>	.68	.26	.36	.23	-.09	-.05	<b>.66</b>	.68

Table 1: Factor structure of the Berlin Mood Questionnaire dependent on presentation form. Kaiser-Meyer-Olkin each .95, Bartlett-Index  $p < 0.01$ , bold: factor load  $> .50$  and proportion of its load on its communality at least 20% higher than the communality proportion of its next higher load on a further factor.

In addition, further application fields for computer assisted diagnostics will present themselves in future that principally cannot be realized using paper versions. In this context we mention the possibility of using the internet or, for instance, using so-called "computer adaptive tests" (CAT) which realize the selection and number of succeeding items to the response behavior appearing in the items already answered [WBK00, CC00]. In this way, items appropriate for the patient are selected on one hand, thus improving precision of measurement, and on the other hand fewer items are needed for measurement altogether. Acceptance of these types of procedures is likely to be far greater than that of conventional ones.

## Bibliography

- [Ame86] American Psychological Association (APA). Guidelines for computer-based tests and interpretations, 1986. Washington DC.
- [BHK93] P Bader, K Hofman, and K-D Kubinger. Zur Brauchbarkeit der Normen von Papier-Bleistift-Tests für die Computer-Vorgabe: Ein Experiment am Beispiel des Gießen-Tests.



*Zeitschrift für Differentielle und Diagnostische Psychologie*, 14(2):129–135, 1993.

- [CC00] D Cella and CH Chang. A discussion of item response theory and its applications in health status assessment. *Med Care*, 38((9 Suppl):II):66–72, 2000.
- [Fah94] J Fahrenberg. Ambulantes Assessment. Computerunterstützte Datenerfassung unter Alltagsbedingungen. *Diagnostica*, 40(3):195–216, 1994.
- [FB89] CC French and JG Beaumont. A computerized form of the Eysenck Personality Questionnaire: A clinical study. *Personality and Individual Differences*, 10:1027–1032, 1989.
- [Fow85] RD Fowler. Landmarks in computer-assisted psychological assessment. *Journal of Consulting and Clinical Psychology*, 53:748–759, 1985.
- [Fra97] GH Franke. Effekte der computerunterstützten Anwendung bei klinisch-psychodiagnostischen Selbstbeurteilungsinventaren unter besonderer Berücksichtigung der Itemreihenfolge. Habilitationsschrift, Institut für Medizinische Psychologie Universität-Gesamthochschule Essen, 1997.
- [FRA01] H Fliege, M Rose, P Arck, S Levenstein, and BF Klapp. Validierung des "Perceived Stress Questionnaire" (PSQ) an einer deutschen Stichprobe. *Diagnostic*, 47(3):142–152, 2001.
- [FRC02] H Fliege, M Rose, L Cotta, M Bullinger, and BF Klapp. Der Fragebogen Alltagsleben: Restrukturierung und klinische Validierung. *Zeitschrift für Medizinische Psychologie*, 2002. in print.
- [Git90] I Gitzinger. Akzeptanz der Darbietung eines Test auf dem Personalcomputer von stationären Patient/-innen. *Psychotherapie, Psychosomatik und medizinische Psychologie*, 40:143–145, 1990.
- [HG86] RR Hart and MA Goldstein. Computer assisted psychological assessment. *Computers in Human Services*, 1:69–75, 1986.
- [HK93] M Hörhold and BF Klapp. Testung der Invarianz und der Hierarchie eines mehrdimensionalen Stimmungsmodells auf der Basis von Zweipunkterhebungen an Patienten- und Studentenstichproben. *Z med Psychol*, 1:27–35, 1993.
- [Hon88] LM Honaker. The equivalency of computerized and conventional MMPI administration: A critical review. *Clinical Psychological Review*, 8:561–577, 1988.
- [Jäg90] RS Jäger. Computerdiagnostik – Ein Überblick. *Diagnostica*, 36:96–114, 1990.
- [KF91] K-D Kubinger and M-G Farkas. Zur Brauchbarkeit der Normen von Papier-Bleistift-Tests für die Computer-Vorgabe: Ein Experiment am Beispiel der SPM von Raven als kritischer Beitrag. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 12(4):257–266, 1991.
- [KM68] B Kleinmuntz and RS McLean. Computers in behavioral science: Diagnostic interviewing by digital computer. *Behavioral Science*, 13:75–80, 1968.
- [Kub95] Kubinger. *Einführung in die psychologische Diagnostik*. Psychologie Verlags Union, Weinheim, 1995.
- [LR94] GA Lienert and U Raatz. *Testaufbau und Testanalyse*. Beltz, Psychologie-Verl.-Union, Weinheim, 1994.
- [RFH00] M Rose, H Fliege, M Hildebrandt, J Körber, P Arck, A Dignass, and BF Klapp. Validierung der deutschsprachigen Version des "Short Inflammatory Bowel Disease Questionnaire" (SIBDQ). *Z Gastroenterol*, 38:277–286, 2000.

- [RHS99] M Rose, V Heß, G Scholler, E Brähler, and BF Klapp. Mobile computergestützte psychometrische Diagnostik: Oekonomische Vorteile und Ergebnisse zur Teststabilität. *Psychother Psychosom med Psychol*, 49:202–207, 1999.
- [SH93] P Schwenkmezger and P Hank. Papier-Bleistift- versus computerunterstützte Darbietung von State-Trait-Fragebögen: eine Äquivalenzüberprüfung. *Diagnostica*, 39(3):189–210, 1993.
- [SRD00] F Schoeneich, M Rose, G Danzer, P Thier, C Weber, and BF Klapp. Narzißmusinventar-90 (NI-90). Empiriegeleitete Itemreduktion und Identifikation veränderungssensitiver Items des Narzißmusinventars zur Messung selbstregulativer Parameter. *Psychother Psychosom med Psychol*, 50:396–405, 2000.
- [Tes86] Testkuratorium der Förderation Deutscher Psychologenvereinigungen. Richtlinien für den Einsatz elektronischer Datenbearbeitung in der psychologischen Diagnostik. *Psychologische Rundschau*, 37:163–165, 1986.
- [WBK00] JE Ware Jr, JB Bjorner, and M Kosinski. Practical implications of item response theory and computerized adaptive testing: a brief summary of ongoing studies of widely used headache impact scales. *Med Care*, 38((9 Suppl):II):73–82, 2000.