

# Wo es beim HaPerT noch hapert – Erfahrungsbericht über einen Handheld-unterstützten Performanztest (HaPerT) für ärztliche Grundfertigkeiten

M. Schmidts, M. Kemmerling, S. Hönigschnabl, M. Lischka  
Institut für medizinische Aus- und Weiterbildung, Medizinische Fakultät, Universität Wien

## Zusammenfassung

Beim HaPerT-System werden OSCE-Checklisten durch Handheld-Computer ersetzt. Am Ende jeder Prüfung werden die Prüfungsergebnisse aller teilnehmenden Kandidaten auf einen Desktop-Rechner überspielt und sofort ausgewertet. Dies ermöglicht individuelles formatives und summatives Feedback an die Kandidaten, selektives Gruppenfeedback sowie die Itemanalyse zur formativen Evaluation des Kurses. Das HaPerT-System wird seit dem Sommersemester 2000 im Lernstudio für ärztliche Grundfertigkeiten eingesetzt. Bei den Prüfern beobachteten wir keine Umstellungsschwierigkeiten. Der Aufwand in Administration und Logistik ist nicht angestiegen. Durch die einfachere und vollständigere Datenerfassung und Auswertung besteht hingegen erstmals die Möglichkeit, den Kurs zu monitoren und im Sinne des Qualitätszirkels zu verbessern.

## Schlüsselwörter

Strukturierte Beobachtung · Feedback · formative Evaluation · OSCE-Logistik · computerunterstütztes Prüfen

## Experiences with the Use of a Handheld Based Performance Test (HaPerT) in a Practical Clinical Skills Course

The HaPerT-System uses handheld computers to replace checklists in an OSCE. At the end of the examination the data of all participating candidates are transferred to a desktop computer. This process allows immediate data evaluation, like individual feedback printouts showing markings for each station, group feedback selected by item weight or item frequency or item analysis (for formative course evaluation). The HaPerT-System is used at our clinical skills center since March 2000. Observers easily familiarized with the new marking approach. With equal administrative and logistic efforts we now are enabled to collect, process and interpret the performance of our candidates in an easier and more elaborated way, which should further improve the quality and outcome of our teaching.

## Key words

Structured observation · feedback · formative evaluation · OSCE logistics · computer based assessment

## Problemstellung und Ziel

Seit acht Jahren wird vom Institut für medizinische Aus- und Weiterbildung ein freiwilliges, von Peer-Tutoren gestütztes Blockpraktikum zum Erlernen praktischer ärztlicher Grundfertigkeiten an Phantomen angeboten [1]. Am Ende der einwöchigen Lehrveranstaltung wird zur Leistungskontrolle der Studenten ein Performanztest (3 Stationen, 3 Examinatoren, strukturierte Beobachtung mittels Checklisten) in Anlehnung an den OSCE (objective structured clinical examination) durchgeführt [2]. Neben der daraus resultierenden summativen Beurteilung könnten die Beobachtungsergebnisse auch zur Qualitätskontrolle der Lehrveranstaltung (z. B. Feedback auf Unterricht der Peer-Tutoren) und zum direkten formativen Feedback an die Studierenden aufbereitet werden. Dieser „Datenmehrwert“ setzt allerdings eine aufwändige Itemanalyse der Checklistenenergebnisse voraus, welche bisher aus ökonomischen Gründen nicht durchgeführt wurde (händische Eingabe der Checklistenitems in eine Datenbank oder Anschaffung teurer Beleglesesysteme). Da die OSCE-Philosophie ein direktes Feedback an die Kandidaten während der strukturierten Beobachtung ausschließt, beschränkte sich unsere Rückmeldung an die Studierenden (nicht zuletzt auch aufgrund der hohen Kursteilnehmerzahlen) in der Regel auf die Mitteilung einer Note.

Primäres Ziel des HaPerT-Projekts war es, die Datenerhebung zu ökonomisieren und gleichzeitig die Datenauswertung und das formative Feedback an Studierende, Kursleiter und Kursplaner im Sinne eines Qualitätszirkels (Abb. 1) weiterzuentwickeln [3,4]. Dies schien uns insbesondere hinsichtlich der Umwandlung des Kurses in eine Pflichtlehrveranstaltung im Rahmen des neuen Medizin-Curriculum Wien [5] (und der damit verbundenen Steigerung der Teilnehmerzahlen) obligatorisch.

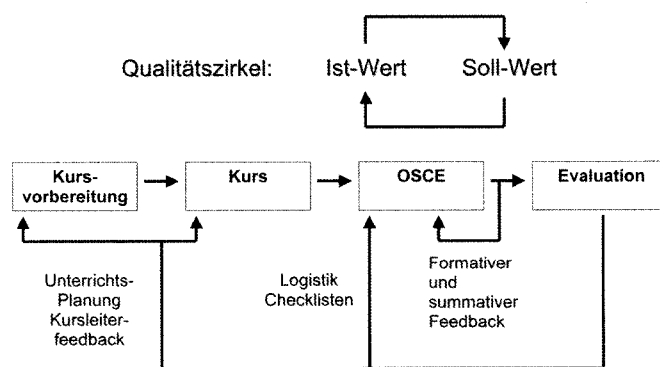


Abb. 1 Formative Kursevaluation.

## Handhelds ersetzen Checklisten

Seit dem Sommersemester 2000 hat die Datenerfassung mittels Handheld-Computer die Checklisten in unserem Kurs ersetzt [6]. Handhelds sind klein, mobil und werden mittels Stifteingabe auf einem berührungsempfindlichen Bildschirm gesteuert.

Die geringe Bildschirmgröße machte es notwendig, die Struktur der Checklisten auf mehrere Eingabemaschinen aufzuteilen und die Dateneingabe zu vereinfachen. Statt jeden Handlungsschritt „ab-

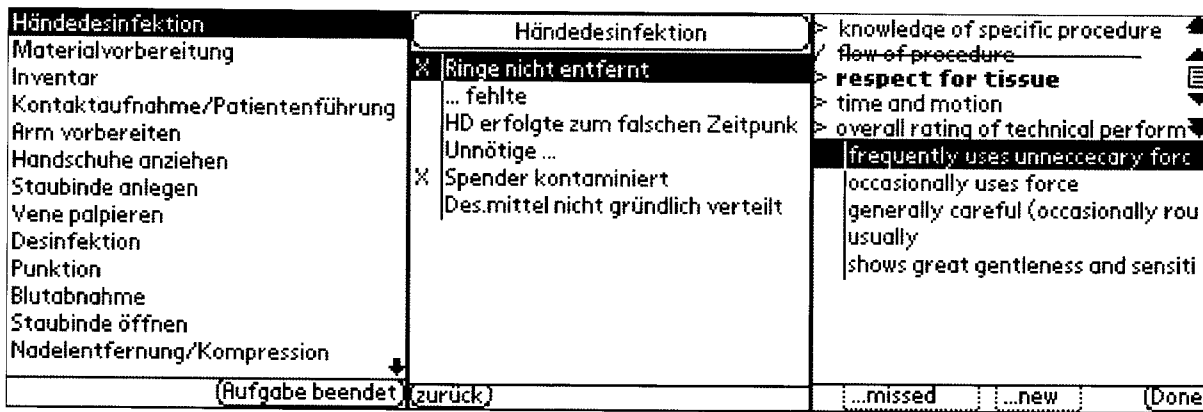


Abb. 2a–c Neugestaltung der Checklisten.

zuhaken“, beschränken wir unsere Aufzeichnungen nun auf inkomplette, fehlerbehaftete oder fehlende Handlungen.

Abb. 2a u. b illustrieren die Aufbereitung der Checkliste „Blutabnahme“. Der Handlungsablauf ist in Handlungsschritte aufgeteilt (Abb. 2a). Jeder Handlungsschritt enthält wiederum Fehleritems als nähere Definitionen des beobachteten Fehlverhaltens (Abb. 2b). Beobachter können auch selbst neue Items oder freie Kommentare eingeben.

Neben Checklisten ist auch der Einsatz von Beurteilungsskalen („rating scales“) mit sich ausschließenden Items möglich (Abb. 2c).

### OSCE-Logistik

Der Handheld-Erfinder Jeff Hawkins prägte den Satz: „Think of the handheld as a ‚tentacle‘ reaching back to the desktop“ [7]. Gemäß diesem Leitgedanken fungieren unsere Eingabegeräte als Eingabemasken von Daten, die auf einem PC verwaltet werden. Abb. 3 zeigt die logistische Integration des HaPerT-Systems in einen OSCE:

Vor jeder Prüfung werden Checklistenitems und Kandidaten auf ein Handheld-Gerät überspielt („synchronisiert“) und von diesem mittels „Schneeballprinzip“ an alle weiteren im Test verwendeten Geräte propagiert („gebeamt“). Am Ende des Tests werden die Kleincomputer wieder eingesammelt und die Prüfungsergebnisse nun in umgekehrter Reihenfolge auf einen einzelnen Handheld zurückübertragen. Dieses Gerät resynchronisiert seine Daten dann mit dem Desktoprechner zur Auswertung.

Im unmittelbaren Anschluss an die Prüfung erhalten die Kandidaten einen Feedbackausdruck, der ihre individuellen Fehler zusammenfasst. Den Kursleitern werden die häufigsten und gefährlichsten Fehlleistungen rückgemeldet.

Tab. 1 fasst alle Auswertungen, welche die PC-Datenbank derzeit leistet, zusammen.

Tab. 1 Derzeit vom HaPerT-System geleistete Auswertungsmechanismen und deren möglicher Einsatz in der Lehrveranstaltung

Auswertung	möglicher Einsatz
individuelle Auflistung aller Fehler eines Kandidaten	Einzelfeedback am Kursende
Zusammenfassung der Fehleritems nach Häufigkeit und Gewichtung pro Kurs	Gruppenfeedback am Kursende Kursleiterfeedback
Zusammenfassung der Fehleritems nach Häufigkeit und Gewichtung pro Semester	Kursplanung Tutoreneinschulung
Prüferschwierigkeiten	Tutoreneinschulung
Stationsschwierigkeiten	Prüfungszusammenstellung Kursplanung
Prüfungsdauer an einzelnen Stationen	Prüfungszusammenstellung Checklistendesign
freie Prüferkommentare	Studentenfeedback Checklistenupdate

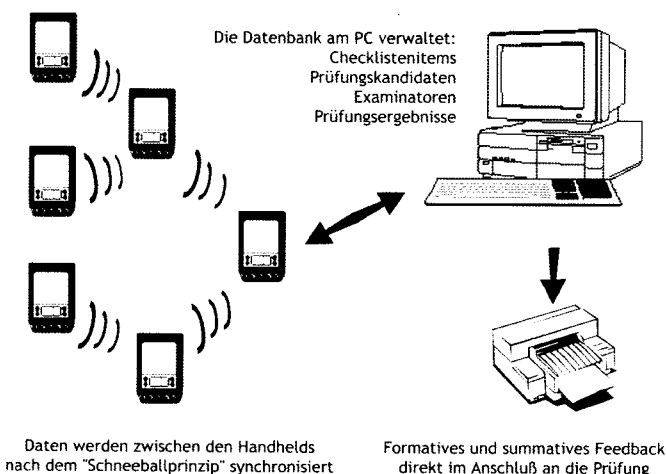


Abb. 3 Logistik des HaPerT-Systems.

## Wahl der Plattform

Als Handhelds benutzen wir die Modelle Palm IIIx sowie Vx. Die mobile Checklistenapplikation wurde mit der Entwicklungsumgebung Satellite-Forms (Hersteller Pumatech) realisiert. Am stationären Desktop (Microsoft Windows) kommt eine mittels Visual Basic for Applications programmierte Access-Datenbank zum Einsatz.

Mit dem Palm IIIx wurde bei der Planung im Jahr 1999 der populärste Handheld-PC gewählt. Ebenso handelte es sich bei der eingesetzten Entwicklungsumgebung Satellite-Forms um die damals einzig verfügbare Skript-Programmiersprache für das PalmOS, die einen vertretbaren Programmieraufwand rechtfertigte. Weitere Argumente für die Plattform waren der im Vergleich zum Konkurrenzprodukt PocketPC (Microsoft) günstigere Anschaffungspreis der Geräte (Einsatz in der Lehre), höhere Systemstabilität und die längere Batterielaufzeit (ausfallssicherer Einsatz bei Prüfungen). Trotz des derzeit rasanten Vormarsches der PocketPCs sind alle vor drei Jahren getroffenen Annahmen – noch – gültig. Dieses Beispiel illustriert treffend die Schwierigkeiten der Auswahl einer zukunftssicheren Betriebssystemplattform (Glück trotz Verstand?).

## Bedienungsfreundlichkeit

Die initiale Dateneingabe in die Handhelds erfordert mehrere Navigationsschritte (Abb. 2) und ist damit unübersichtlicher und auch etwas umständlicher als eine lineare Checkliste. Weiters zwingt uns die kleine Bildschirmgröße manchmal Handlungsbeschreibungen der Checkliste stark zu kürzen („Instr. ‚Kompr. Punkt.stelle‘ falsch“ bedeutet beispielsweise „Instruktion an den Patienten ‚Komprimieren Sie die Punktionsstelle‘ war falsch“).

Die Benutzung des Systems setzt in vermehrtem Maße einen gewissen „Expertenstatus“ des Beobachters voraus, da die Handlung nicht mehr Schritt für Schritt abgehakt wird, sondern die „Fehler“ aktiv erkannt und selektiv aufgezeichnet werden müssen.

Den unseren Kurs unterstützenden Peer-Tutoren, die neben ihrer Unterrichtstätigkeit auch bei der Prüfung mitwirken, bereitete die Umstellung auf Handheld-Computer trotz oben genannter Bedenken keine Probleme. Unsere bisherigen Erfahrungen haben im Gegenteil gezeigt, dass eine Einschulungsphase von ca. 15 Minuten völlig ausreicht, um sich mit der Bedienung der Geräte und der neuen Struktur der Checklisten vertraut zu machen.

Für die Eingabe sowie das Update der Checklisten in die PC-Datenbank, die Datenverwaltung der Kandidaten und Examinatoren, die Auswertung der Testergebnisse und die Koordination der Prüfung wird ferner ein speziell eingeschulter Systemadministrator benötigt.

## „Work in Progress“

Das HaPerT-System wurde seit der Einführung vor zwei Jahren kontinuierlich im Funktionsumfang erweitert. Vor allem die Möglichkeit, während der Prüfung auch freie Kommentare in das System einzugeben, wurde von unseren Evaluatoren stark genutzt und ließ die Itemdatenbank stetig anwachsen. Das zunehmende (auch internationale) Interesse an der HaPerT-Software machte weitere Adaptationen der Software und der Checklisten-Designs notwendig. Derzeit sind wir bestrebt, das Benutzerinterface der PC-Datenbank zu vereinfachen.

Aufgrund dieser Entwicklungs- und Umstellungsprozesse, sowie der Bindung der Systemadministration an eine Person, kam es in den vergangenen zwei Jahren immer wieder zu Nutzungspausen des HaPerT-Systems, die leider eine vollständige Erhebung der Semesterleistung aller Kandidaten (und damit eine systematische formative Kursevaluation) bisher verhinderten.

## Aufwand

Sieht man vom Entwicklungsaufwand ab, halten sich Routinetätigkeiten des HaPerT-Systems (Checklistenpflege, Kandidatenverwaltung, Vorbereitung, Abwicklung und Auswertung der Prüfung) mit dem klassischen Checklisten-System insgesamt die Waage. Initialer Zusatzaufwand kostet die Implementierung des Systems, der Ankauf der Handhelds (ca. € 350,-/Stück) und die Einschulung eines Systemadministrators, der allerdings als Examinator am Check beteiligt sein kann.

## Feedback

Trotz der Vielfalt von Variationen wird der OSCE seinem Wesen nach als summative Abschlussprüfung eingesetzt. Die Quasi-Echtzeit-Datenerfassung und -Datenverarbeitung des HaPerT-Systems bietet den für uns entscheidenden Vorteil des sofortigen formativen Einzel- und Gruppenfeedbacks. Der individualisierte Feedbackausdruck reflektiert die Leistung in der Prüfung viel nachvollziehbarer als die Note. Die in der Prüfung aufgetretenen häufigsten, gefährlichsten oder systematisch auftretenden Fehler können in einer Gruppenabschlussrunde strukturiert rückgemeldet werden. Damit erhält der Kursleiter gleichzeitig wertvolle Hinweise über mögliche Auslassungen oder Missverständnisse im Unterricht.

Aufgrund unseres an Fehlhandlungen orientierten Checklisten-Designs erhalten die Prüfungskandidaten derzeit ausschließlich negatives Feedback (eine Auflistung aller Fehler, die gemacht wurden). Ein balanciertes Feedback wird derzeit entwickelt.

## Formative Evaluation

Kursleiterfeedback und formative Kursevaluation im Sinne des Qualitätszirkels bilden das noch nicht systematisch umgesetzte Kernstück des HaPerT-Projekts (Abb. 1).

Tab. 2 Gesamtanalyse der Fehleritemhäufigkeit für die Station „Blutabnahme“ im Sommersemester 2001 (26 strukturierte Beobachtungen)

Handlungsschritte	Fehleritems	n
Verabschiedung	¥ Instr. „1 min fest draufdrücken“	6
Staubinde öffnen	Staubinde nicht geöffnet	4
Nadelentfernung	Nadel zu spät entsorgt	4
Inventar	¥ Ersatznadel	3
Materialentsorgung	Staubinde beim Patient vergessen	3
Punktion	eine Fehlpunktion	3
Staubinde öffnen	Staubinde zu spät geöffnet	3
Blutabnahme	zu wenig Blut im Röhrchen	3

Tab. 2 zeigt exemplarisch Auswertungsergebnisse für das Sommersemester 2001 für die Station „Blutabnahme“. Die Daten weisen auf die problematischen Punkte „Patienteninstruktion“, „Staubinde öffnen“ und „Nadelentfernung“ hin.

Ziel ist es, am Semesterende eine ausführliche Fehleritemanalyse (Darstellung der Itemhäufigkeit wahlweise global, getrennt nach OSCE-Stationen, getrennt nach Kursen oder nach Prüfern) durchzuführen und diese Informationen für die Kursplanung einzusetzen. Einsatzmöglichkeiten wären:

- die Schulung der Unterrichtenden,
- die Reflexion der Lernziele,
- die Feststellung von Ausbildungsdefiziten und Ausbildungsdifferenzen sowie
- Hinweise auf Elemente des „hidden curriculum“.

### Ausblick

Im Bereich des Peer-Teaching liegt die Qualitätssicherung des Unterrichts auf der Hand. Frühere Untersuchungen haben allerdings ergeben, dass im Bereich der praktischen ärztlichen Grundfertigkeiten ganz allgemein ein sehr starker Strukturierungs- und Standardisierungsbedarf besteht [1]. Im neuen Medizincurriculum wird unser Kurs vermehrt von klinisch tätigen Universitätslehrern unterrichtet werden. Diese werden vermutlich in ihrer Durchführungspraxis verschiedener Grundfertigkeiten erheblich voneinander differieren, also sozusagen das „hidden curriculum“ der Fertigkeiten repräsentieren. Die strukturierte Beobachtung mittels Checkliste deckt solche unkommunizierten Lernziele und Handlungsdiskrepanzen zwischen Fachkollegen unweigerlich auf, entweder durch die Kritik des Universitätslehrers an der Checkliste (Kritik am kommunizierten Lernziel) oder durch die Beurteilung eines Prüfungskandidaten als Handlungsträger des Unterrichts eines Fachkollegen (Kritik am unkommunizierten Lernziel). Das HaPerT-Projekt könnte derart als objektivierbare Kommunikationsplattform zur gemeinsamen Definition von Lernzielen ausgeweitet werden, quasi als Neuinterpretation der studentenzentrierten Feststellung von George Bordages Aussage „Assessment drives the curriculum“ [8]. Aus dem Qualitätszirkel würde so eine Qualitätsspirale: Experten kritisieren, korrigieren und gewichten Checklistenitems, entweder vor der Prüfung oder durch freie Kommentareingabe wäh-

rend der Prüfung. Diese Items müssen sich dann in der Nutzung behaupten, d.h. von den Kollegen auch angekreuzt werden („Itemdarwinismus“). Häufig auftretende Fehler wiederum gäben einen Hinweis auf differierende Lehrmeinungen, die es abzuklären gälte.

Wir möchten uns an dieser Stelle herzlich bei Dr. Thomas Link, Monika Himmelbauer und Eva M. Zechmeister für die Durchsicht des Manuskripts bedanken.

### Literatur

- <sup>1</sup> Schmidts M, Beran H, Lischka M. Patient versus Kunststoffpuppe – Profitieren Studenten mit Famulaturerfahrung vom systematischen Training klinischer Grundfertigkeiten. Zeitschrift für Hochschuldidaktik 1998; 22: 75–82
- <sup>2</sup> Harden R, Gleeson F. Assessment of clinical competence using an objective structured clinical examination (OSCE). Medical Education 1979; 1: 41–54
- <sup>3</sup> Falk-Ytter Y. Prüfungsstrukturen für eine erfolgreiche Studienreform. Medizinische Ausbildung 2001; 18: 133–137
- <sup>4</sup> Eitel F, Tesche A. Learning medical competence – physician heal themselves! Zeitschrift für Hochschuldidaktik 1999; 23: 41–50
- <sup>5</sup> Studienplan des Diplomstudiums Humanmedizin. URL: [www.univie.ac.at/mcw/studienplan/index.html](http://www.univie.ac.at/mcw/studienplan/index.html)
- <sup>6</sup> Schmidts M. OSCE logistics-handheld computers replace checklists and provide automated feedback. Medical Education 2000; 34: 957–958
- <sup>7</sup> Rhodes N, McKeehan J. Palm Programming – the Developers Guide. Sebastopol: O'Reilly, 1998; 6
- <sup>8</sup> Bordage G. Assessment drives the curriculum: What-Why-How to assess student learning. Abstractband 4. Europäischer Kongress „Qualität der Lehre in der Medizin“, Bern 4.–6. September 1997: 11

**Korrespondenzadresse:** Dr. med. Michael Schmidts · Institut für med. Aus- und Weiterbildung · Abteilung für Forschung und Entwicklung · Universität Wien · Postfach 10 · 1097 Wien · Österreich · E-mail: michael.schmidts@akh-wien.ac.at