

Panoramic Displays – Bewertung von Benutzerschnittstellen im Kontext von Mehrfachtigkeit

Armin Eichinger, Johannes Kellerer, Peter Sandl und Alf Zimmer

Schlüsselwörter: Großflächendisplay, Evaluationskonzept, Touchscreen, Belastungskontext

Zusammenfassung

Usability-Untersuchungen berücksichtigen häufig nicht den Kontext, in dem das untersuchte Artefakt eingesetzt wird. In hoch belastenden Umgebungen wird der Kontext vorwiegend von der Art der informatorischen Belastung, der Höhe der Belastung und von der Verteilung der Aufmerksamkeit auf Aufgaben bei Mehrfachttigkeiten beeinflusst.

Untersuchungsergebnisse im Rahmen einer vergleichender Bewertung von Touchscreen und Trackball bei Zeigeaufgaben weisen signifikante Einflüsse dieser Größen auf die Bedienleistung nach. Die differenzierten Ergebnisse rechtfertigen den aufwändigen Evaluationsansatz.

Summary

Usability studies often neglect the context in which the artefact under study is being used. In high workload environments, context is mainly influenced by the quality of information processing demands, by the quantity of these demands, and by the allocation of attentional resources between multiple tasks.

Results of a comparative evaluation of a touchscreen and a trackball used in aiming tasks reveal significant influences of these variables on aiming performance. The differential results justify the elaborate and time-consuming evaluation approach.

Einleitung

Studien zur Untersuchung der Usability verschiedener Benutzerschnittstellen werden häufig so konzipiert, dass eine Bewertung innerhalb eines experimentellen Durchgangs anhand eines Leistungskriteriums erfolgen kann. Die entscheidende Rolle der Umgebung für die Interaktion wird experimentell nicht aufgegriffen. Für die Untersuchung der Interaktion mit komplexen Systemen in dynamischen und hoch belastenden Umgebungen, wie dem Cockpit von Kampfflugzeugen, greift dieser Evaluationsansatz daher zu kurz. Zwar erkennen viele Autoren die Bedeutung der mentalen Belastung und erheben bei ihren Studien die subjektive Beanspruchung oder Workload als abhängige Variable (z. B. Grandt et al., 2003). Das experimentelle Vorgehen entspricht trotzdem häufig dem, das Rouse (1985) wie folgt kritisiert: “The tendency is to study problems that are scaled down to fit the methods being investigated. Simplifying assumptions are made in order to allow rigorous formulations. Problem-specific contingencies are eliminated in order to promote experimental control“ (S. 3).

Im Folgenden wird ein Ansatz zur experimentellen Evaluation von Benutzerschnittstellen vorgestellt, der den differenzierten Einfluss der Arbeitsumgebung auf die Bedienleistung explizit berücksichtigt und variiert. Der beschriebene Evaluationsansatz umfasst nicht nur die Untersuchung der Usability-Maße, sondern erstreckt sich auch auf die experimentelle Analyse von Aspekten der Wahrnehmbarkeit und des situativen Verständnisses, die im Konstrukt Situationsbewusstsein zusammengefasst werden. Der aktuelle Stand des Projekts erlaubt die Vorstellung erster Ergebnisse zur Usability-Untersuchung.

Evaluation

Hintergrund

Die Versuche fanden im Rahmen des Projekts *Panoramic Display* statt. Ziel des Projekts ist es zu untersuchen, ob ein Großflächendisplay mit Touchscreenfunktion als *Head Down Display* verwendet werden kann. Gegenstand der vorliegenden Untersuchung ist die vergleichende Bewertung von Touchscreen und Trackball als Eingabemedien für die Interaktion mit einem Großflächendisplay.

Konzept

Der vorliegende Beitrag beschreibt einen Ansatz zur Bewertung von Benutzerschnittstellen, der dem differenzierten Einfluss des Arbeitsumfelds auf die Qualität der Interaktion Rechnung trägt. Leistungsdaten werden bei gleichzeitiger Ausführung von zusätzlichen Aufgaben erhoben. Der Belastungskontext wird dabei hinsichtlich seiner (1) Qualität, (2) Quantität und (3) Gewichtung manipuliert. Diese Attribute identifiziert Wickens (2008) als entscheidende Bestimmgrößen mentaler Beanspruchung.

1. Die **Qualität** des Belastungskontextes wird über die Art der Zusatzaufgabe beeinflusst. Drei Zusatzaufgaben decken die für die Flugaufgabe dominierenden Mechanismen der Informationsverarbeitung isoliert ab, wie sie Wickens (2003, 2006) in seinem klassischen Modell beschreibt: Anforderungen der *visuellen Wahrnehmung*, werden durch eine visuelle Suchaufgabe abgebildet. Das klassische Paradigma der Suche im Kurzzeitgedächtnis erzeugt Anforderungen der *kognitive Verarbeitung*. Eine isoliert manuell-motorische Aufgabe deckt den Aspekt der *Handlungsausführung* ab. Entsprechend den Resultaten einer vorab durchgeführten Aufgabenanalyse wird eine komplexe Zusatzaufgabe verwendet, die den Belastungskontext im Cockpit nicht für isolierte Informationsverarbeitungsstufen, sondern möglichst realitätsnah über alle Verarbeitungsinstanzen abbilden soll. Diese Aufgabe setzt sich aus einer kompensatorischen Tracking- sowie einer verbalen umgekehrten Wiederholaufgabe, einer sog. *Readback*-Aufgabe zusammen.

2. Die **Quantität** der Belastung wird über zwei unterschiedliche Schwierigkeitsgrade der Zusatzaufgaben manipuliert.

3. Um dem Umstand Rechnung zu tragen, dass der Aufmerksamkeitsschwerpunkt nicht notwendig auf die Zeigeaufgabe am Großflächendisplay gerichtet ist, wird die **Priorisierung** der Aufgaben beeinflusst. So werden die Zusatzaufgaben hinsichtlich ihrer Bedeutung durch entsprechende Instruktionen der Probanden als Nebenaufgaben, als Hauptaufgaben oder als gleichwertig der Hauptaufgabe gewichtet und entsprechend bearbeitet.

Als Leistungsindikatoren wurden Maße der Genauigkeit und Geschwindigkeit erhoben, als zusätzliches Effizienzmaß die subjektive Beanspruchung der Teilnehmer über Fragebögen.

Im Rahmen klassischer Zeigeaufgaben waren einzeln erscheinende Ziele möglichst schnell und genau anzuwählen. Nach Anwahl wurden die Ziele ausgeblendet. Das Erscheinen der Ziele war zeitgesteuert und erfolgte nicht direkt nach Anwahl.

Methode

Die Versuche wurden in einem Cockpit-Mockup der Abteilung *Human Factors Engineering* der EADS in Manching durchgeführt. Display, Bediengeräte und Teilnehmer wurden entsprechend den Gegebenheiten im Cockpit positioniert. Zur Teilnahme erklärten sich neun Testpiloten der EADS im Manching im Alter von 28 bis 52 Jahren bereit. Alle Probanden nahmen freiwillig an den Untersuchungen teil.

Als unabhängige Variable wurden Eingabeelement (Touchscreen vs. Trackball), Art der Zusatzaufgabe (visuell, kognitiv, motorisch, komplex), Schwierigkeit der Zusatzaufgabe (leicht,

schwer) und Priorisierung (Schwerpunkt Hauptaufgabe, Schwerpunkt Nebenaufgabe, gleich) in allen Kombinationen untersucht.

Als abhängige Variable wurden Zeit- und Fehlermaße für die Zeigeaufgaben erhoben. Die Leistungsmaße für die standardisierten Zusatzaufgaben waren Anzahl der Einzeldurchgänge für die visuelle und motorische sowie Reaktionszeit für die kognitive Zusatzaufgabe. Bei der komplexen Zusatzaufgabe wurden für die Tracking-Komponente die über die Zeit gemittelte Abweichung von der Bildschirmmitte (der sog. *Root Mean Squared Error*, RMSE) und für die Readback-Aufgabe der Grad der Übereinstimmung mit der korrekten Reihenfolge ermittelt.

Die einzelnen Versuchsdurchgänge dauerten 90 Sekunden. Die Reihenfolge ihrer Bearbeitung war randomisiert. Jeder Teilnehmer bearbeitete alle Versuchsbedingungen.

Ergebnisse

Die Leistungsmaße der Zeigeaufgaben wurden in einem vollständig faktoriellen Design erhoben: Zusatzaufgabe x Eingabemedium x Schwierigkeit x Priorisierung. Das Design für die Untersuchung der Zusatzaufgabe weist den Faktor „Zusatzaufgabe“ nicht auf: Eingabemedium x Schwierigkeit x Priorisierung. Exemplarisch für weitere Effektivitäts- und Effizienzmaße zur Erhebung der Zeigeleistung wird die varianzanalytische Untersuchung der Bearbeitungsdauer dargestellt. Zusammenfassend werden die Aufgabenleistungen in Abbildung 1 dargestellt.

Zeigeaufgabe

Es konnten signifikante Haupteffekte für die Art der Zusatzaufgabe ($F_{3,24} = 26.50$, $p < .001$), das Eingabemedium ($F_{1,8} = 164.320$, $p < .001$), die Schwierigkeit der Zusatzaufgabe ($F_{1,8} = 32.35$, $p < .001$) sowie der Priorisierung der beiden Aufgaben festgestellt werden ($F_{2,16} = 35.23$, $p < .001$).

Post Hoc-Vergleiche nach Bonferroni ergaben eine signifikant bessere Bedienleistung bei gleichzeitiger kognitiver Belastung verglichen mit den anderen Zusatzaufgaben. Weitere Einzelvergleiche der Zusatzaufgaben wurden nicht signifikant. Die Zeigeaufgabe wurde signifikant schneller unter leichter als unter schwerer Zusatzbelastung ausgeführt. Die Ausführung war mit dem Touchscreen signifikant schneller als mit dem Trackball. Der Einfluss der Priorisierung zeigte eine signifikante Steigerung Bediengeschwindigkeit entsprechend der Verteilung der Aufmerksamkeit.

Zusatzaufgaben

Die folgende Tabelle zeigt Teststatistiken und Irrtumswahrscheinlichkeiten der untersuchten Haupteffekte für alle Zusatzaufgaben:

	visuell	kognitiv	motorisch	Tracking	Readback
Eingabemedium	$F_{1,8} = 1.01$ $p = .345$	$F_{1,8} = 0.01$ $p = .948$	$F_{1,8} = 9.24$ $p = .016$	$F_{1,8} = 21.10$ $p = .002$	$F_{1,8} = 3.08$ $p = .117$
Schwierigkeit	$F_{1,8} = 474.87$ $p < .001$	$F_{1,8} = 1.95$ $p = .200$	$F_{1,8} = 491.80$ $p < .001$	$F_{1,8} = 2.52$ $p = .151$	$F_{1,8} = 11.17$ $p = .010$
Priorisierung	$F_{2,16} = 5.55$ $p = .015$	$F_{2,16} = 1.18$ $p = .331$	$F_{2,16} = 5.16$ $p = .019$	$F_{2,16} = 5.88$ $p = .012$	$F_{2,16} = 0.71$ $p = .508$

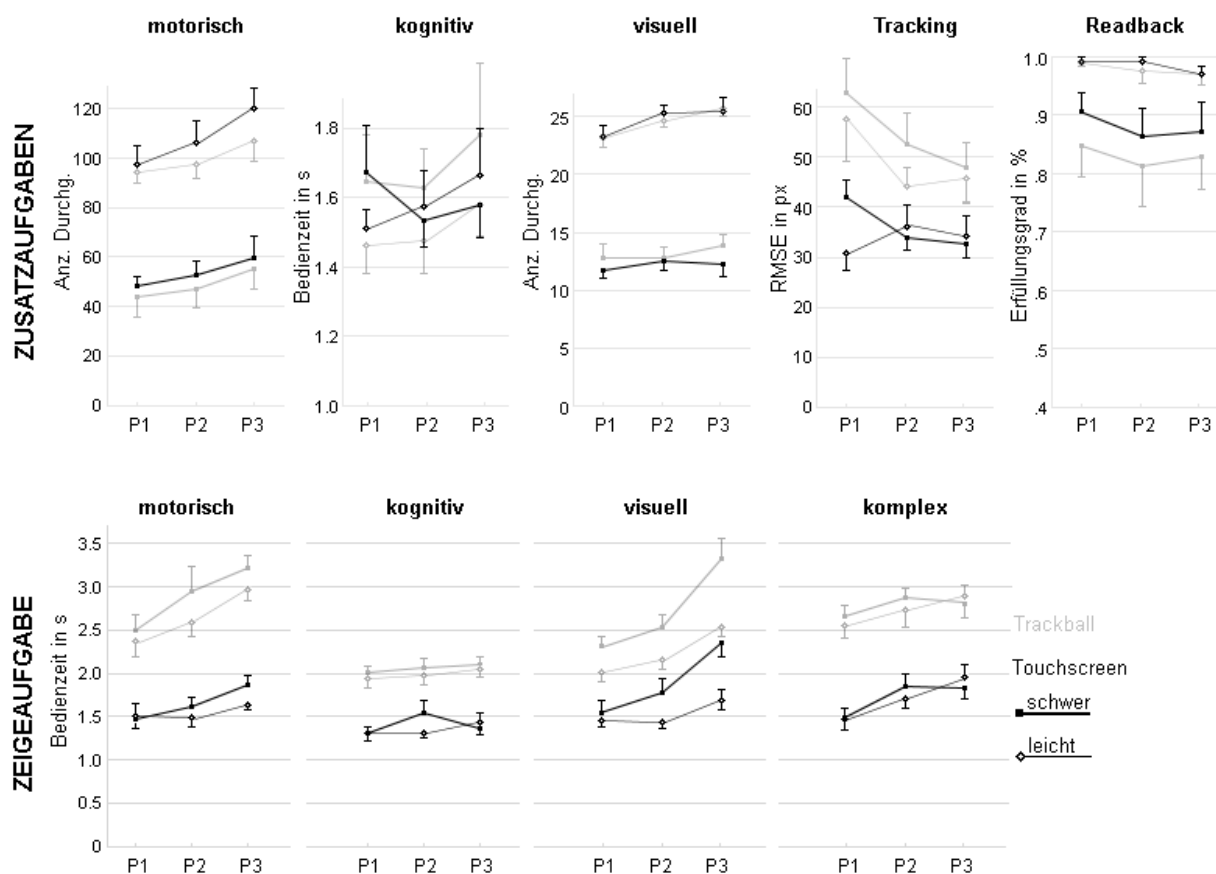


Abb.1: Leistung in den Zusatzaufgaben (*oben*) und Bearbeitungszeit bei der Zeigeaufgabe (*unten*); die Fehlerbalken entsprechen einem Standardfehler. P1: Priorität Zeigeaufgabe, P2: gleiche Gewichtung, P3: Priorität Zusatzaufgabe.

Diskussion

Die differenzierte Natur der Ergebnisse legt nahe, dass jedes der manipulierten Attribute des Belastungskontextes für die Erhebung klassischer Usability-Maße in hoch belastenden und dynamischen Arbeitsumgebungen untersucht werden sollte. Der beschriebene Ansatz kann analog für Leistungskriterien des Situationsbewusstseins eingesetzt werden. Der Aufwand für diese Art der bewertenden Untersuchung ist hoch; Umfang und Qualität der gewonnenen Erkenntnisse, von denen hier nur ein kleiner Ausschnitt dargestellt werden konnte, tragen sicher zur Amortisierung dieser Anstrengungen bei.

Literatur

- Grandt, M. and Pfendler, C. and Mooshage, O. (2003). Empirical Comparison of Five Input Devices for Anti-Air Warfare Operators. In: Proceedings of the 8th ICCRTS on "Information Age Transformation", 17-19 June 2003, Washington DC.
- Rouse, W. B. (1985). On better mousetraps and basic research: Getting the applied world to the laboratory door. *IEEE Transactions on Systems, Man, and Cybernetics*, 2-8.
- Wickens, C. D. (2008). Multiple Resources and Mental Workload, *Human Factors*, 50(3), 449-455.
- Wickens, C. D. (2003). Aviation displays. In P. Tsang & M. Vidulich (Eds.), *Principles and practice of aviation psychology* (pp. 147-199). Mahwah, NJ: Lawrence Erlbaum.
- Wickens, C.D., and Carswell, C.M (2006). *Human Information Processing*. In G. Salvendy (Ed.) *Handbook of Human Factors*. N.Y.: John Wiley & Sons.