

**Autor(en):** Prümper, J.

**Titel:** Der Benutzungsfragebogen ISONORM 9241/10: Ergebnisse Zur Reliabilität und Validität

**Jahr:** 1997

**Laufende Nr. im Literaturverzeichnis:** ISONORM 9241/10 [4]

**Schlagwörter (3-5):** ISONORM 9241/10, Reliabilität, Validität, Normierung

**Kurzbeschreibung:** Das Paper evaluiert den von Prümper, J. & Anft, M. (1993) entwickelten Fragebogen ISONORM 9241/10, mit dem die Usability von Software nach den ISO-Kriterien zur benutzerfreundlichen Dialoggestaltung (ISO 9241-10) bewertet werden kann. Dazu liegen Beurteilungen von 1265 Benutzern vor, die insgesamt 178 Softwareprogramme evaluiert haben. Es werden hierbei Aussagen zur Reliabilität, Validität und zur Normierung des Fragebogens gemacht. Die Reliabilität wurde mittels Retestrelabilität untersucht, wobei bei einer Messzeitpunktdifferenz von 6,7 Monaten und Werten von .59 bis .68 ( $p < .001$ ) von einer konstanten Merkmalerfassung durch den ISONORM 9241/10 ausgegangen werden kann. Des Weiteren wurde die innere Homogenität der Items mit Cronbach's alpha ( $\alpha < .81$ ,  $p < .01$ ) untersucht. Die angegebenen Werte sprechen für eine hohe Reliabilität. Für die Erfassung der Validität wurden die kriterienbezogene Validität und die Konstruktvalidität herangezogen. Für die innere kriterienbezogene Validität wurde der ISONORM 9241/10 mit dem QUIS ( $r = .73$ ,  $p < .001$ ) und dem BBD ( $r = .71$ ,  $p < .001$ ) verglichen. Es zeigt sich, dass der ISONORM zu ähnlichen Ergebnissen gelangt wie andere, valide Benutzungsfragebögen. Um die äußere kriterienbezogene Validität zu erfassen, wurde ein Vergleich mit dem EVADIS II als objektives Kriterium durchgeführt ( $r = .59$ ,  $p < .01$ ). Es zeigt sich, dass der Fragebogen hohe Ähnlichkeiten mit der Expertenevaluation aufweist. Für die Konstruktvalidität wurde untersucht, inwiefern Verbesserungen einer Software durch den Fragebogen erfasst werden. Dabei wird deutlich, dass Unterschiede zwischen den Versionen mit dem ISONORM 9241/10 messbar sind. Durch die Verwendung von 7 Skalen lassen sich darüber hinaus differenzierte Aussagen über Stärken und Schwächen der Versionen treffen. Um die Beurteilung des Fragebogens mit anderen Verfahren vergleichen zu können, wurde eine Normierung an 41 Programmen durchgeführt.

Als besondere Stärke der Untersuchung sind die Größe der Stichprobe und die Verwendung verschiedener Programme zur Evaluation des Fragebogens zu nennen. Daher kann von einer breiten Verwendbarkeit des Fragebogens ausgegangen werden, die der Autor postuliert. Das Paper ist gut verständlich und wichtige Informationen für die Nachvollziehbarkeit der Untersuchung (z.B. Angaben zum Alphaniveau, zu Substichproben und verwendeter Instrumente) sind gegeben. Insgesamt kann die testtheoretische Güte als sehr positiv beurteilt werden, was für ein robustes, vielseitig anwendbares Verfahren spricht.

**Erstellt von:** Sophie Neef, Matrikelnr.: 346936