

Modellierung verschiedener Meßverfahren zur Erfassung der Ordnungsschwelle

H. Meister, H. Klüser, M. Walger, H. von Wedel
Universitäts-HNO-Klinik Köln

Zusammenfassung

In der hier präsentierten Studie werden mehrere in der Literatur beschriebene Verfahren zur Erfassung der Ordnungsschwelle mittels Monte-Carlo-Simulationen untersucht. Dabei kann festgestellt werden, daß die Methoden bei gleicher zugrundeliegender psychometrischer Funktion und damit identischer Perzeptionsleistung deutlich voneinander abweichende Werte produzieren. Auch hinsichtlich Genauigkeit und Effizienz ergeben sich Unterschiede. Dies kann ein wesentlicher Grund für die unterschiedlichen Angaben von Ordnungsschwellenwerten sein.

Grundsätzlich kann die Ordnungsschwelle bei geeigneter Parameterwahl und entsprechender Auswertung des Zielwertes mit einfach zu realisierenden adaptiven Meßverfahren zufriedenstellend schnell und genau bestimmt werden.

1. Einleitung

Die Ordnungsschwelle (OS) beschreibt den minimalen Zeitabstand (hier: "interaural time differences", ITD), bei dem zwei Sinnesreize mit einer bestimmten Wahrscheinlichkeit p_r hinsichtlich ihres Auftretens erkannt und in eine Reihenfolge gebracht werden. Die OS wird als Maß für das zeitliche Auflösungsvermögen häufig im Zusammenhang mit der Diagnostik und Therapie zentral-auditiver Verarbeitungsstörungen (ZAVST) genannt, hinsichtlich ihrer Ergebnisse und deren Interpretierbarkeit jedoch kontrovers diskutiert.

Eine mögliche Ursache hierfür können die deutlich unterschiedlichen psychophysikalischen Meßverfahren sein, die im Rahmen der Erfassung der OS zur Anwendung kommen. Generell werden verschiedene adaptive Meßverfahren und Festwertmethoden eingesetzt, die größtenteils nicht evaluiert sind und z.T. eher intuitiv entwickelt wurden.

Im Rahmen der hier vorgestellten Studie werden einige in der Literatur beschriebene Methoden systematisch untersucht.

2. Material und Methoden

Die im folgenden aufgeführten Meßverfahren zur Erfassung der Ordnungsschwelle wurden mit einem rechnergestützten Modell überprüft, wobei Monte-Carlo-Simulationen zum Einsatz kamen. Hierbei wurde eine typische psychometrische Funktion (PF) eines Normalhörenden zugrunde gelegt, die zuvor empirisch bestimmt wurde. Da Methode 1 (s.u.) den Zielwert nach etwa 75 trials - einem durchaus praxisbezogenen Wert - erreichte, fand dieser Wert als Maximaldarbietung der Vergleichbarkeit halber auch bei allen anderen Verfahren Verwendung.

Methode 1 (Ilmberger et al. 1986): "Increasing method of limits", Darbietung von jeweils 8 Reizpaaren, beginnend bei ITD=100 ms. Werden mindestens 6 von 8 Stimuli hinsichtlich ihrer Reihenfolge richtig beurteilt, wird die ITD um 10 ms verringert. Werden weniger als 6 Reize richtig beurteilt, wird der vorhergehende Wert als OS

definiert. Die Messungen werden einmal wiederholt und gemittelt. Konvergiert auf $p_r=75\%$.

Methode 2 (Warnke 1995): Gewichtetes adaptives Verfahren mit einer Schrittweite (SW) von 1 ms. Beginnend bei 100 ms wird die ITD bei jeder richtigen Beurteilung um eine SW verringert, bei jeder falschen um 4 SW erhöht. Der so erreichte Bestwert wird als OS definiert. Ein Abbruchkriterium ist nicht definiert und wird individuell vom Untersucher festgelegt. Konvergiert auf $p_r=80\%$.

Methoden 3a und b (Meister et al. 1998): a) Adaptives 2-Schrittverfahren, Startwert ITD=100 ms. SW=10 ms. Bei zwei aufeinanderfolgenden richtigen Beurteilungen wird die ITD um 1 SW verringert, bei jeder falschen um 1 SW erhöht. Konvergiert auf $p_r=70,7\%$.

b): Adaptives 3-Schrittverfahren. Startwert ITD= 100 ms. Anfangs-SW=20 ms, wird bei jedem Wendepunkt halbiert bis SW minimal 5 ms. Konvergiert auf $p_r=79,4\%$. Die OS wird bei beiden Verfahren durch Mitteln der präsentierten ITDs vom 4. bis zum 14. Wendepunkt bestimmt.

Methoden 4a und b (Nickisch 1999): a) Gewichtetes adaptives Verfahren beginnend bei ITD=100 ms. SW beträgt 2 ms. Verringerung um eine SW bei richtiger Beurteilung, Erhöhung um vier SW bei falscher Antwort. Konvergiert auf $p_r=80\%$.

b) wie a) jedoch Verringerung und Erhöhung um je 5 ms. Konvergiert auf $p_r=50\%$.

Bei beiden Methoden dient die adaptive Steuerung lediglich zur Plazierung der Werte. Die OS wird dann als der kleinste Wert definiert, bei dem 80% der für diese ITD dargebotenen Reize richtig beurteilt wurden.

3. Ergebnisse

Abbildung 1 zeigt die Änderung des Gesamtfehlers der Ordnungsschwellenschätzung in Abhängigkeit von der Anzahl der Darbietungen für die adaptiven Methoden 2-4 und gibt damit einen Anhaltspunkt für die Effizienz (Abnahme des Fehlers mit zunehmender Anzahl der trials) der Verfahren. Die unterschiedlichen Auswertungen sind hierbei nicht berücksichtigt.

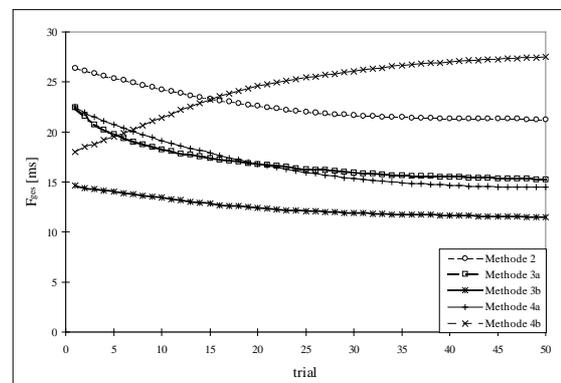


Abb.1: Abhängigkeit des Gesamtfehlers ($F_{ges} = \sqrt{F_{sys}^2 + F_{zuf}^2}$) von der Anzahl der Darbietungen (trials) ab dem 4. Wendepunkt.

Auffällig ist das Ansteigen des Fehlers für Methode 4b. Dies ist dadurch zu begründen, daß sich der

Schätzwert mit zunehmender Anzahl der Darbietungen immer weiter vom zugrunde gelegten Sollwert x_{80} entfernt, da die Methode an sich auf die Ratewahrscheinlichkeit konvergiert. Für alle anderen Verfahren nimmt der Gesamtfehler immer weiter ab. Hierbei ist deutlich zu erkennen, daß Methode 4a aufgrund der größeren Schrittweite effizienter als Methode 2 ist. Den insgesamt kleinsten Fehler weist Methode 3b (3-Schritt-Verfahren) auf.

Tabelle 1 zeigt den Zielwert (X_{soll}) der Ordnungsschwelle, den "bias" (systematischer Fehler F_{sys} =Abweichung von Zielwert) und die Streuung SD (zufälliger Fehler F_{zuf}) der OS-Werte der verschiedenen Methoden unter Berücksichtigung der beschriebenen Auswertungen und gibt damit ein Maß für die Genauigkeit der Verfahren.

Meth.	1	2	3a	3b	4a	4b
x_{soll}	29,5	37	24	36	37	37
bias	5,7	9,5	2,0	0,7	7,2	-13,8
SD	14,7	5,5	7,6	8,4	9,9	8,5

Tab. 1: OS-Zielwert (x_{soll}) sowie systematischer (bias) und zufälliger Fehler (SD) der Methoden, basierend auf dem Endergebnis nach 75 trials. Angaben in ms.

Hierbei zeigen Methoden 3a und b den geringsten, 2 und 4b den höchsten bias. Methode 2 zeigt die geringste, Verfahren 1 die bei weitem höchste Standardabweichung. Grundsätzlich ist hierbei auch die Steigung der psychometrischen Funktion zu berücksichtigen, die für größere p_r generell flacher verläuft und damit für diese Verfahren tendenziell höhere Standardabweichungen erzeugt. Dies spielt auch beim Vergleich der Methoden 3a und 3b eine Rolle.

5. Zusammenfassung

Die dargestellten Ergebnisse zeigen, daß die in der Literatur beschriebenen Verfahren zur Messung der Ordnungsschwelle deutlich unterschiedliche Ergebnisse produzieren. Dies ist nicht verwunderlich, da aus wahrscheinlichkeitstheoretischen Gründen unterschiedliche Richtig-Antwortwahrscheinlichkeiten p_r und damit verschiedene x_{soll} zugrundeliegen. Daher ist ein direkter Vergleich von OS-Werten ohne Angabe von p_r nicht statthaft.

Auch hinsichtlich ihrer Genauigkeit unterscheiden sich die Verfahren. Die größte Variabilität und damit die geringste Genauigkeit weist Methode 1 auf, wobei sich die Streuung aufgrund der vorgeschlagenen Mittelung über zwei Messungen theoretisch noch um den Faktor 1,41 reduziert. Bei der Vorgabe, von nur 8 Stimuli 6 richtig beurteilen zu müssen, damit die jeweilige ITD akzeptiert wird, spielt bei einem 2AFC-Verfahren die Ratewahrscheinlichkeit eine große Rolle. Auch der Binomialtest weist für diese Konfiguration keine Signifikanz aus ($p=0,289$).

Die kleinste Streuung zeigt Methode 2, was durch die extrem kleine Schrittweite von 1 ms zu begründen ist. Allerdings ist diese Wahl ineffizient, da der theoretische Zielwert erst nach sehr vielen Darbietungen erreicht wird, was zu einem großen bias und zu einem hohen Gesamtfehler führt. Insbesondere bei Untersuchungen von Kindern, bei denen Motivation und Vigilanz eine entscheidende Rolle spielen, kann das langsame Konvergenzverhalten problematisch sein.

Deutlich unter dem theoretischen Wert produziert Methode 4b Ergebnisse. Diese konvergiert auf $p_r=50\%$, was bei einem 2AFC-Verfahren der Ratewahrscheinlichkeit entspricht. Da die Methode entsprechend die Werte im Bereich der Ratewahrscheinlichkeit plazierte, wird die tatsächliche OS erheblich unterschätzt, wenn bei der Auswertung ein 80%-Kriterium angelegt wird.

Die Untersuchungen zeigen weiterhin, daß mit sehr einfachen, bekannten Verfahren (z.B. Methoden 3a und b), durchaus effektiv und ausreichend genau gemessen werden kann. Dies gilt bei sinnvoller Parameterwahl (Startwert und Schrittweite) auch für die leicht zu realisierenden gewichteten adaptiven Verfahren. Allerdings sollte bei allen Methoden ein Abbruchkriterium definiert und nicht vom Untersucher entschieden werden, wie lange überhaupt gemessen wird und wann der (unbekannte) Zielwert annähernd genau geschätzt ist. Dies kann der Einfachheit halber beispielsweise dadurch bewerkstelligt werden, daß eine bestimmte Anzahl von Wendepunkten vorgegeben wird und über die n (gerade Zahl) letzten Wendepunkte der Meßwert durch Mitteln bestimmt wird.

Literatur

- Ilmberger, J.: Auditory excitability cycles in choice reaction time and order threshold. *Naturwissenschaften* 73 (1986), 743-744
- Meister, H., H. Klüser, M. Dück, M. Walger, H. von Wedel: Adaptive Verfahren zur Messung der Ordnungsschwelle. *Zeitschrift für Audiologie* (1998), 110-120
- Nickisch, A.: Ordnungsschwellenwerte im Vor- und Grundschulalter. *Sprache, Stimme, Gehör* 23 (1999): 63-70
- Warnke, F.: *der Takt des Gehirns*. VAK-Verlag (1995), Freiburg

Gefördert von der Jean-Uhrmacher-Stiftung

Dr. Hartmut Meister
email: hartmut.meister@uni-koeln.de