**(Fuchs, 2004)**

* Approach to designing CBM tasks: Design CBM tasks in a way that they represent the curriclum oft hat schoolyear (eg math CBM by Fuchs et al. 1997) 🡪 same amount oft he same problems for every test
* Ist in quop ja auch so gehalten, gibt es eigentlich ein Paper, das quop vorstellt?!
* Measure the skills specified in the curriculum on every test throughout the term for every individual student
* Depict rate of learning (each test score stating how much oft he targeted end of term performance a student has at the moment (Fuchs, Fuchs, Hamlett & Allinder, 1991)) For us that means: Assessing targeted end of term performance broken down into several skills every three weeks systematically
* Research in CBM: Three stages:

1. "Investigation oft he technical features oft he static score (i.e. perormfance level at one point in time)" Fuchs, Fuchs, Hamlett, Walz & Germann, 1993
2. "technical features of slope"; "determine whether increasing CBM scores (or slope) are in fact associated with improvement in overall competence in the academic domain"#+
3. "instructional utility" 🡪 are teachers able to use the info offered by CBM?
   1. Vielleicht als Ausblick für die nächste Untersuchung dieser Testreihe? So "haben jetzt gezeigt, dass Stages 1 und 2 erfüllt sind, man also progress monitoring betreiben kann, aber Nutzen in der Praxis muss untersucht werden"

* Shapiro, Angello and Eckert (2004): Has progress monitoring increased over the past decade? 🡪 Als weitere Punkt, dafür, dass das gut ist?
* Papers 1. Research stage: Ardoin et al., 2004; Clarke & Shinn, 2004; McGlinchey & Hixson, 2004 (letzte Studie: "assessing CBM correlations and decision utility with respect to a highstakes reading test.")
* Hintze and Christ (2004) stage 2: slope!
* "CBM tasks are multidimensional" not "single-skill measurement"

2014 Souvignier Förster Salaschek (Quelle noch hinzufügen)

* "Urteilsgenauigkeit von Lehrkräften unterscheidet sich erheblich", mglw. (Feinberg und Shapiro, 2009)
* "Zum einen müssen die über die Zeit hinweg eingesetzten Testverfahren alle die gleiche Schwierigkeit aufweisen - sie müssen also parallel sein. Zum anderen müssen sich die Aufgaben als sensitiv gegenüber Leistungsveränderungen erweisen und tatsächliche Lerngewinne entsprechend abbilden."
  + - Anforderungen an die Testgüte: "Klassisch" Ojektivität, Reliabilität, Validität (wobei Validität speziell) und "Parallelität der Einzeltests" sowie "Sensitivität der Einzeltests für Leistungsveränderungen"
    - Vielleicht Generell LPA, dann zu den Kriterien kommen, dann sagen Parallelität ist da, jetzt Validität?
* Eine der größten Herausforderungen: Parallele Tests konstruieren
* Gute Praktikailität: 3.4 von 4 von Lehrern
* "In Mathematik wird häufig der Ansatz des curriculum sampling verwendet, bei dem jeder Paralleltest dieselben im Curriculum definierten Teilkompetenzen anhand einer Teilmenge von Aufgaben erfasst (z.B. Additions- und Subtraktionsaufgaben) (Foegen, Jiban & Deno, 2007). Im Gegensatz dazu wird im Bereich des Lesens der Ansatz des robust indicator bevorzugt, bei dem eine Aufgabe verwendet wird, die die Integration verschiedener Teilkompetenzen erfordert (z.B. lautes Lesen für eine Minute) und die hohe Korrelationen zu curricularen Kompetenzen (Leseverständnis) aufweist (Reschly, Busch, Betts, Deno & Long, 2009).
* Hier "Kombination beider Konstruktionsansätze": robuste Indikatoren mit zusätzlicher Orientierung an curricularen Anforderungen
* Für Leseverständnis reicht flüssiges Lesen für eine Minute (passage reading fluency / oral reading fluency (Deno, Mirkin, and Chiang, 1982 (aus Fuchs, 2004)) nicht aus, da es flüssige Leser gibt, die dennoch ein geringes Leseverständnis haben (Lerkkanen, Rasku-Puttonen, Aunola & Nurmi, 2004)

References

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, *33*(2), 188–193.