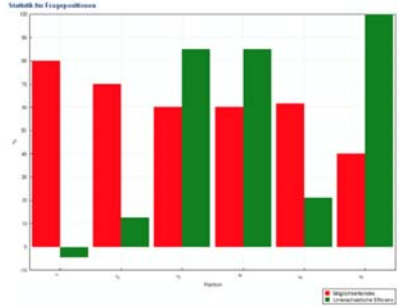
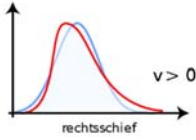
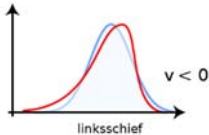
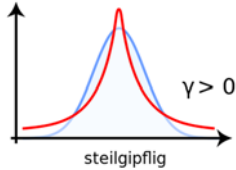
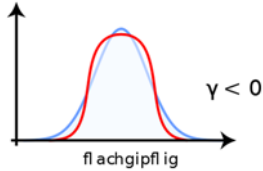


Begriffe (moodle-intern)	Bedeutung	Beispiel
Standardabweichung	wie groß war die Abweichung der einzelnen Bewertungen von der Durchschnittsbewertung	Die richtige Beantwortung einer Quiz-Frage von Studierenden in Moodle besitzt einen Mittelwert von 50%, mit einer Varianz von 44% -> D.h. die durchschnittliche Abweichung der Studierenden von ihrem Mittelwert beträgt 44%
Leichtigkeitsindex	Gibt den Schwierigkeitsgrad einer Frage an Berechnung: $X_p = \text{durchschnittliche Punktzahl bei der Testaufgabe}$ $X_p (\text{min}) = \text{Niedrigste Punktzahl}$ $X_p (\text{max}) = \text{höchste Punktzahl}$	10 Studenten beantworten eine Frage. (10 Punkte sind maximal erreichbar) Durchschnittlich erreichte Punkte: (X_p) 6 Niedrigste erreichte Punktzahl: ($X_p (\text{min})$) 2 Höchste erreichte Punktzahl: ($X_p (\text{max})$) 9 Rechnung: $100 \times ((6-2)/(9-2))$ -> $100 \times (4/7)$ -> $100 \times 0,57$ -> 57 -> umso höher der Index (Maximum ist 100) ist, desto einfacher war die Frage für die Studierenden zu beantworten
Zufällig angenommene Punktezahl	Beschreibt die Punktezahl, welche bei einer richtigen „Rateantwort“ erteilt wird	Die Beantwortung einer Frage wird mit 5 Punkten belohnt, wenn diese jedoch geraten wird (wie genau getestet wird, ob ein Student

		ratet ist hierbei nicht weiter dargelegt worden), werden lediglich 3 Punkte vergeben
Beabsichtigte Gewichtung		wenn eine Frage 3 Punkte für die richtige Antwort bringt und im Test 10 Punkte erreicht werden können, dann ist die beabsichtigte Gewichtung 30%.
Effektive Gewichtung	Symbolisiert das Maß, wie viel die Frage zur Abweichung von der Durchschnittsbewertung im Test insgesamt beigetragen hat Diese sollte im Idealfall nahe an der beabsichtigten Gewichtung liegen	Die erste Frage eines Tests bringt 3 Punkte und es gibt insgesamt 10 Punkte zu erreichen. -> beabsichtigte Gewichtung: 30 % 9 von 10 Studenten beantworten die erste Frage falsch -> effektive Gewichtung höher als die beabsichtigte, da so die Durchschnittsbewertung negativ beeinflusst wird.
Discrimination Index	Ist ein grober Indikator dafür, wie „gut“ eine Frage in Bezug auf den gesamten Test ist -> Teilnehmer, welche in einer „guten“ Frage eine hohe Punktzahl erreichten, sollten auch im Test eine hohe Punktzahl erreichen -> eine „gute“ Frage ist dann eine gute Frage, wenn sowohl der Test, als auch die Frage gut beantwortet wurde Mathematische Berechnung: $DI = (X_{top} - X_{bottom}) / N$	Wenn eine Frage in einem Test zu 80% richtig beantwortet worden ist und der Studierende auch im Test gut bis sehr gut abgeschnitten hat (80%-100%), fällt dieser Indikator hoch aus

	<p>X_{top} = Summe des oberen Drittels der best-erreichten Punktzahlen</p> <p>X_{bottom} = unteres Drittel der best-erreichten Punktzahlen</p> <p>N = Anzahl der Teilnehmer an dem Test</p>	
<p>Unterscheidungseffizienz</p>	<p>Ist ein Indikator dafür, inwieweit eine Frage in einem Test ein Ausreißer darstellt</p>	 <p>->Das Balkendiagramm zeigt für jede einzelne Frage den Möglichkeitsindex (d.h. wie oft prozentual die Frage richtig beantwortet wurde) und die Unterscheidungseffizienz (d.h. ein Maß dafür, ob die Frage als Ausreißer das Gesamtergebnis des Tests verfälscht oder nicht,</p>
<p>Schiefe der Punkteverteilung</p>	<p>Ist eine Kennzahl, die die Art und Richtung einer Asymmetrie einer Wahrscheinlichkeitsverteilung angibt</p>	<p>Schiefe $s=0,443$</p> <p>-> rechtsschief, also negative Veränderung</p> 

		<p>Negativ, weil so weniger Punkte erreicht wurden, als durch die Normalverteilung prognostiziert wurde</p> <p>-> linksschief, also positive Veränderung</p> 
<p>Bewertungsverteilungsgraph</p>	<p>Maßzahl für die Spitzigkeit oder Steilheit einer Wahrscheinlichkeitsverteilung</p> <p>Verteilungen mit geringer Wölbung streuen relativ gleichmäßig, Verteilungen mit hoher Wölbung beinhalten oft einige Ausreißer hinsichtlich der Ergebnisse</p>	<p>Wenn der Wert >0 ist, dann ist der Graph steilgipflig</p>  <p>-> d.h. es gab mehrere Ausreißer im Test, sowohl im positiven, als auch im negativem Sinne</p> <p>Wenn der Wert <0 (bspw. -3,4) ist, dann ist der Graph flachgipflig</p> 

		-> d.h. im Vergleich zur Normalverteilung gibt es nicht viele bzw. bedeutende Ausreißer
Koeffizient interner Konsistenz	bezeichnet das Ausmaß, in dem die Aufgaben bzw. Fragen einer Skala miteinander in Beziehung stehen -> bspw. Bei MC-Fragen sagt dieser Wert aus, inwiefern die Fragen das Gleiche messen	Wenn der Koeffizient einen Wert größer 0,7 annimmt, sind die Fragen in Bezug auf die Konsistenz anwendbar
Fehlerquotient	Ist ein Indikator dafür, wie sehr ein (absolvierter) Test wirklich den Wissensstand der Studierenden widerspiegelt -> Desto geringer der Fehlerquotient ist, umso besser spiegelt der Test den Wissensstand wider	Verhältnis zwischen zufälligen Unterschieden und leistungsbedingten Unterschieden in Bewertungen
Standardfehler	Wird vom Fehlerquotienten abgeleitet und gibt an wie viel zufällige Abweichung in jeder einzelnen Testbewertung steckt	Standardfehler: 10 % Person erreicht im Test: 60 % -> tatsächlicher Wissensstand liegt zwischen 50% und 70%
Möglichkeitsindex	Beschreibt den Anteil der Testversuche, in welcher die Frage richtig beantwortet wurde	Wenn 5 von 10 Fragen im Testversuch richtig beantwortet werden, beträgt der Möglichkeitsindex 50%