

Erhöhung der Validität und der Fehlerrückmeldung digitaler diagnostischer Testaufgaben durch STACK



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Ausgangslage

„An deutschen Hochschulen verzeichnet man seit mehr als einer Dekade den alarmierenden Befund, dass einem Großteil der Studierenden bei Studienbeginn viele mathematische Grundkenntnisse und -fertigkeiten sowie konzeptuelles Verständnis mathematischer Inhalte fehlen.“
(Koepp et al., 2017, S. 1)

Ausgangslage

- Hohe Studienabbruchzahlen in mathematikaffinen Studiengängen, z. B. Mathematik mit 47% (Autorengruppe Bildungsberichterstattung 2014)
- Mathematische Kompensationskurse finden an fast jeder Hochschule statt (Biehler et al. 2014)
- Starke Heterogenität zu verzeichnen (Schiemann 2013)



Notwendigkeit einer Diagnose in Kompensationskursen
(Biehler et al. 2014)



Gliederung

1. **Ausgangslage**
2. STACK als Meilenstein der Testentwicklung
3. Chancen und Grenzen von STACK für die Diagnose
4. Elementarisierendes Testen
5. Ausblick

Gliederung

1. Ausgangslage
2. **STACK als Meilenstein der Testentwicklung**
3. Chancen und Grenzen von STACK für die Diagnose
4. Elementarisierendes Testen
5. Ausblick

Validität

- Keine Ratewahrscheinlichkeit wie beim MC (z.B. Kamps & van Lint, 1975)
- Mehr leere Antworten als über MC mit entsprechender Option (Kallweit, 2016)
- Große Antwortvielfalt (Kallweit, 2016)
- Keine Provokation von Fehlern durch Distraktoren (Nitsch, 2015)
- Intendierte Prozessrichtung valider prüfbar (u.a. Sangwin & Jones, 2017; Katz et al., 1996; Kamps & van Lint, 1975)
- Besonders einschrittige Aufgaben können durch ein Identifizieren statt einem Realisieren der richtigen Antwort gelöst werden (Feld-Caesar, 2017)

Validität

- Keine Ratewahrscheinlichkeit wie beim MC (z.B. Kamps & van Lint, 1975)
- Mehr leere Antworten als über MC mit entsprechender Option (Kallweit, 2016)
- **Hypothese:**
 - Aufgaben mit geringer Komplexität sind im offenen Format empirisch schwieriger als im geschlossenen Format.
(Katz et al., 1996; Kamps & van Lint, 1975)
- Besonders einschrittige Aufgaben können durch ein Identifizieren statt einem Realisieren der richtigen Antwort gelöst werden (Feld-Caesar, 2017)

STACK – Chancen und Grenzen

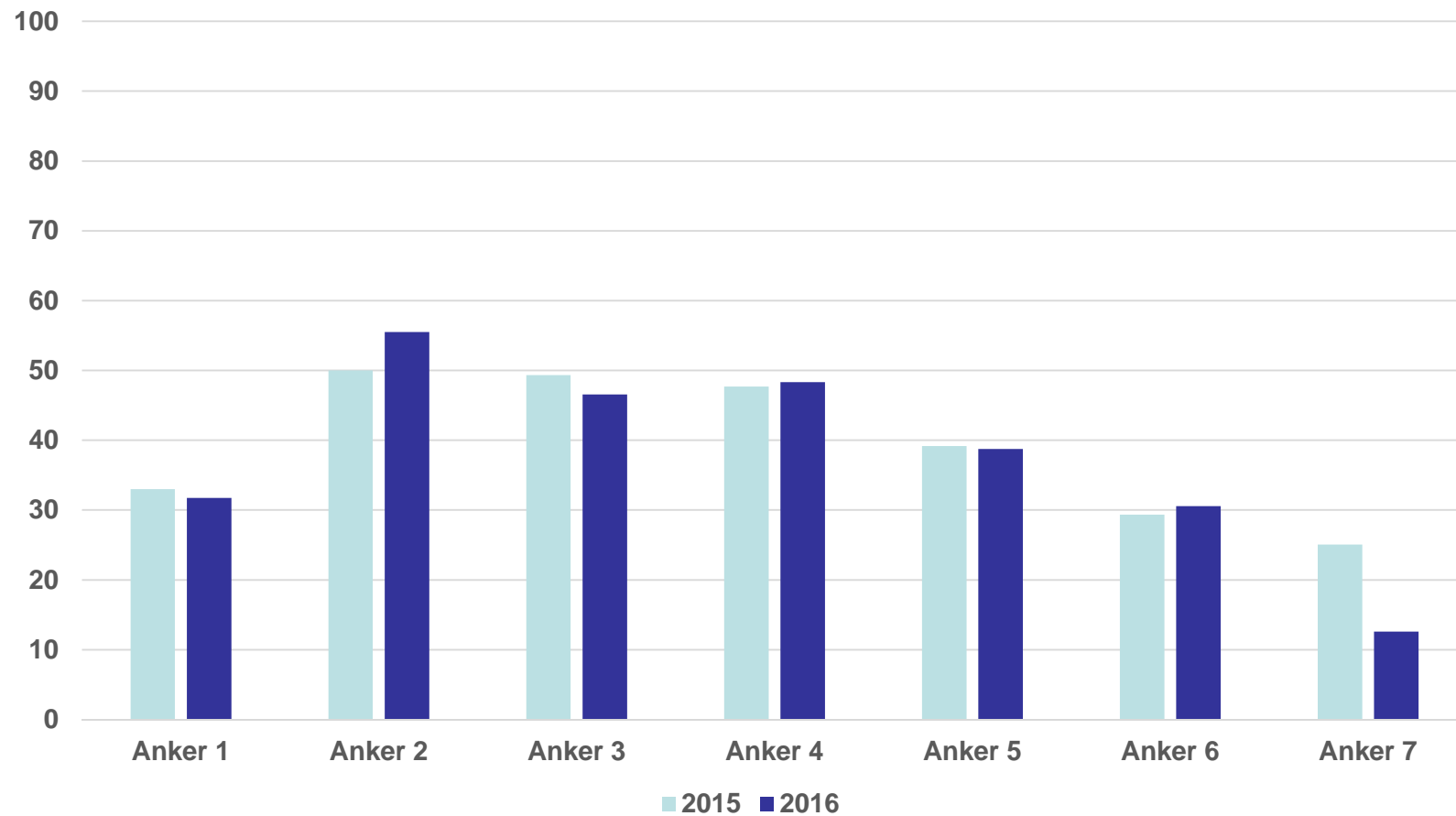


Empirische Untersuchung

- Vergleich von offenen und geschlossenen Format im VEMINT-Eingangstest des WS15/16 (N=868) und WS16/17 (N=627)
- Itemvergleich von 4 weniger komplexen Aufgaben mit Wechsel vom geschlossenen ins offene Antwortformat
- Mann-Whitney-U-Test (Eckstein, 2016)

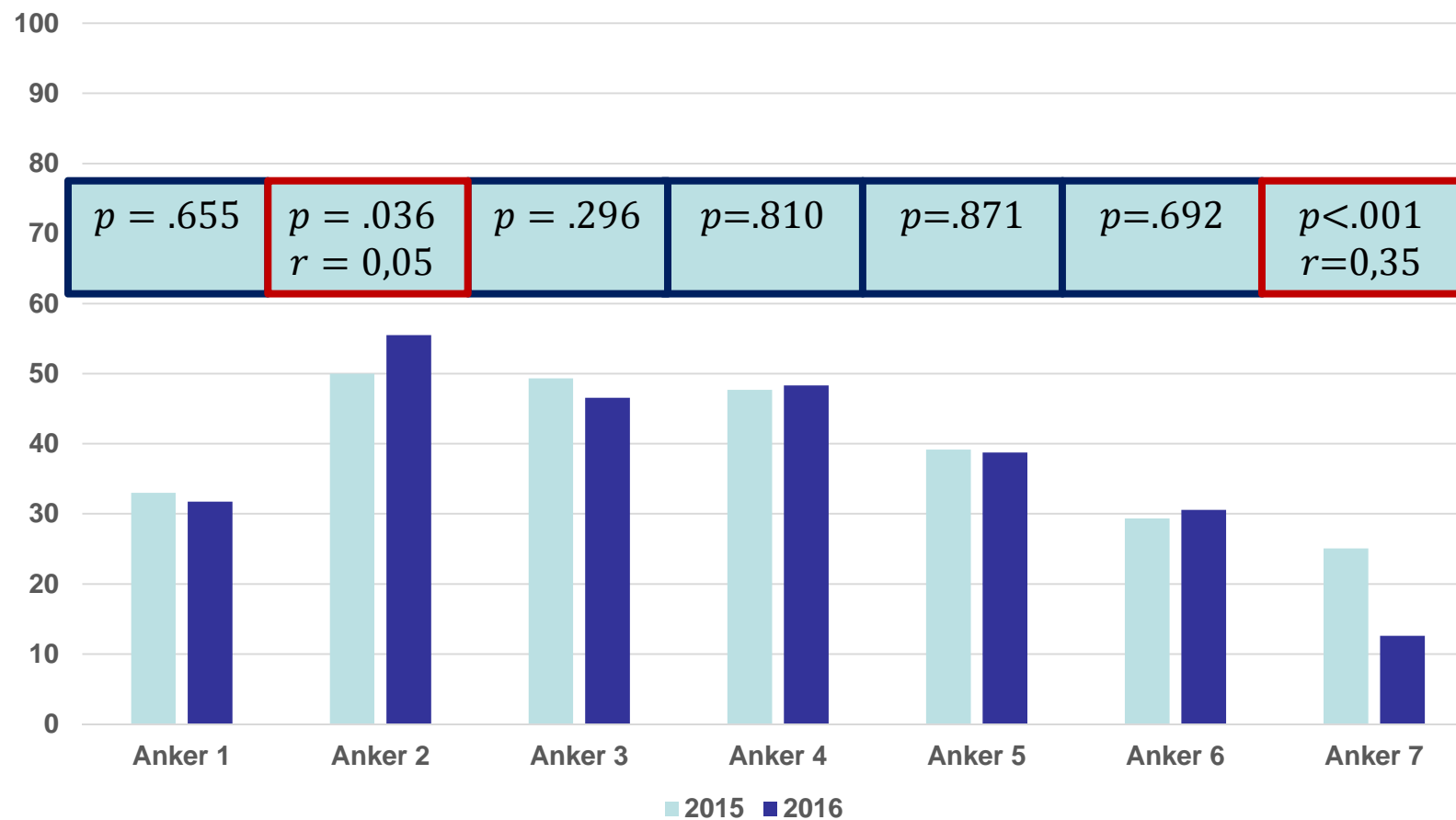
STACK – Chancen und Grenzen

Lösungshäufigkeiten der Ankeritems



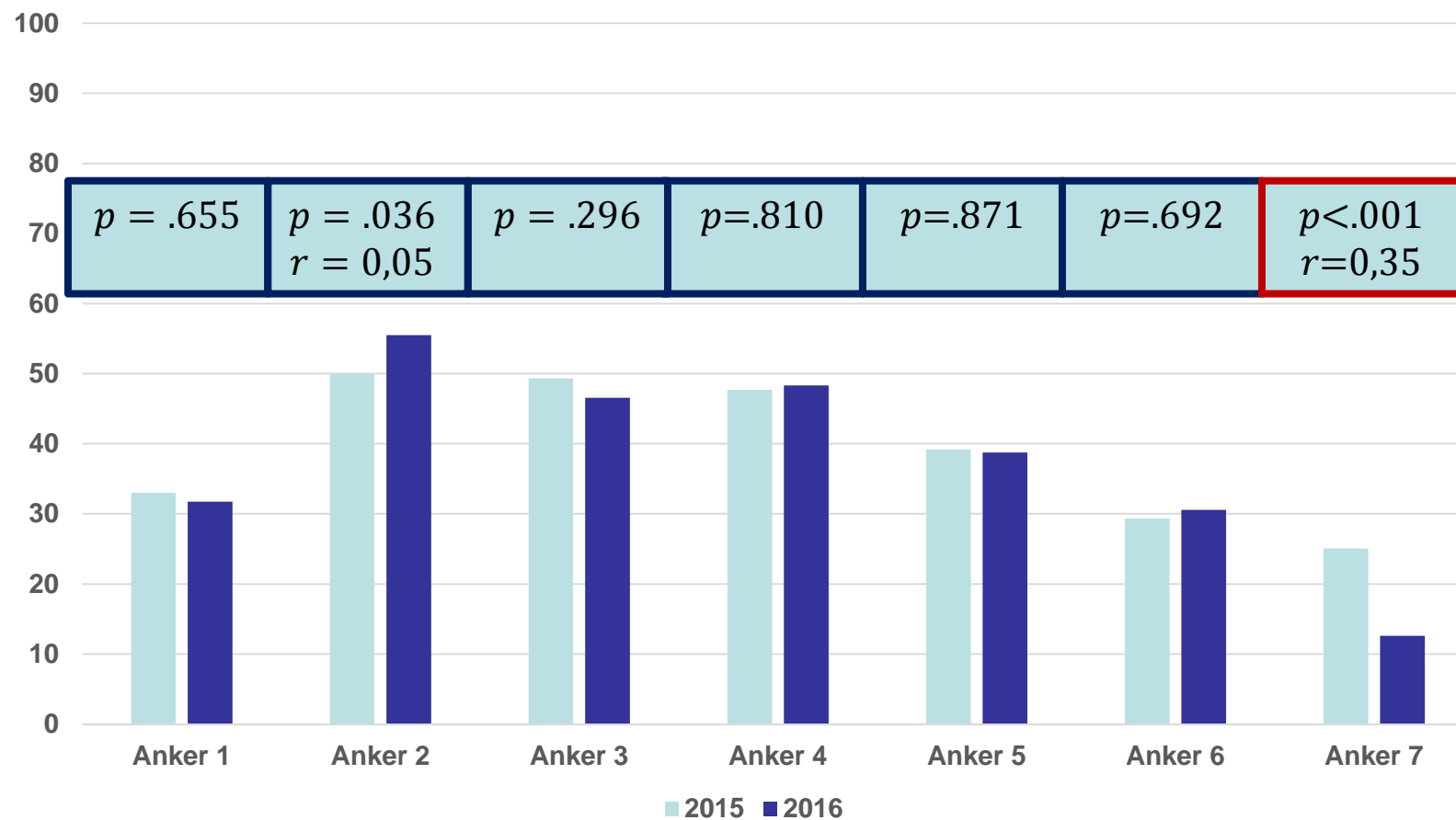
STACK – Chancen und Grenzen

Lösungshäufigkeiten der Ankeritems



STACK – Chancen und Grenzen

Lösungshäufigkeiten der Ankeritems



STACK – Chancen und Grenzen

Lösungshäufigkeiten der Ankeritems

Aufgabe 18

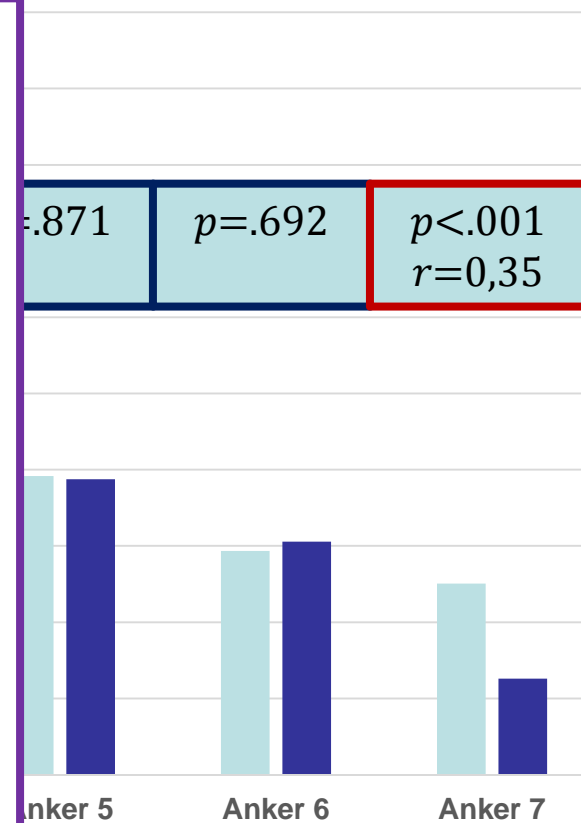
Vereinfachen Sie den Term $\sqrt{\left(\frac{a}{2}\right)^2 + a^2}$ mit $a \geq 0$.

Wählen Sie eine Antwort:

Antwort:

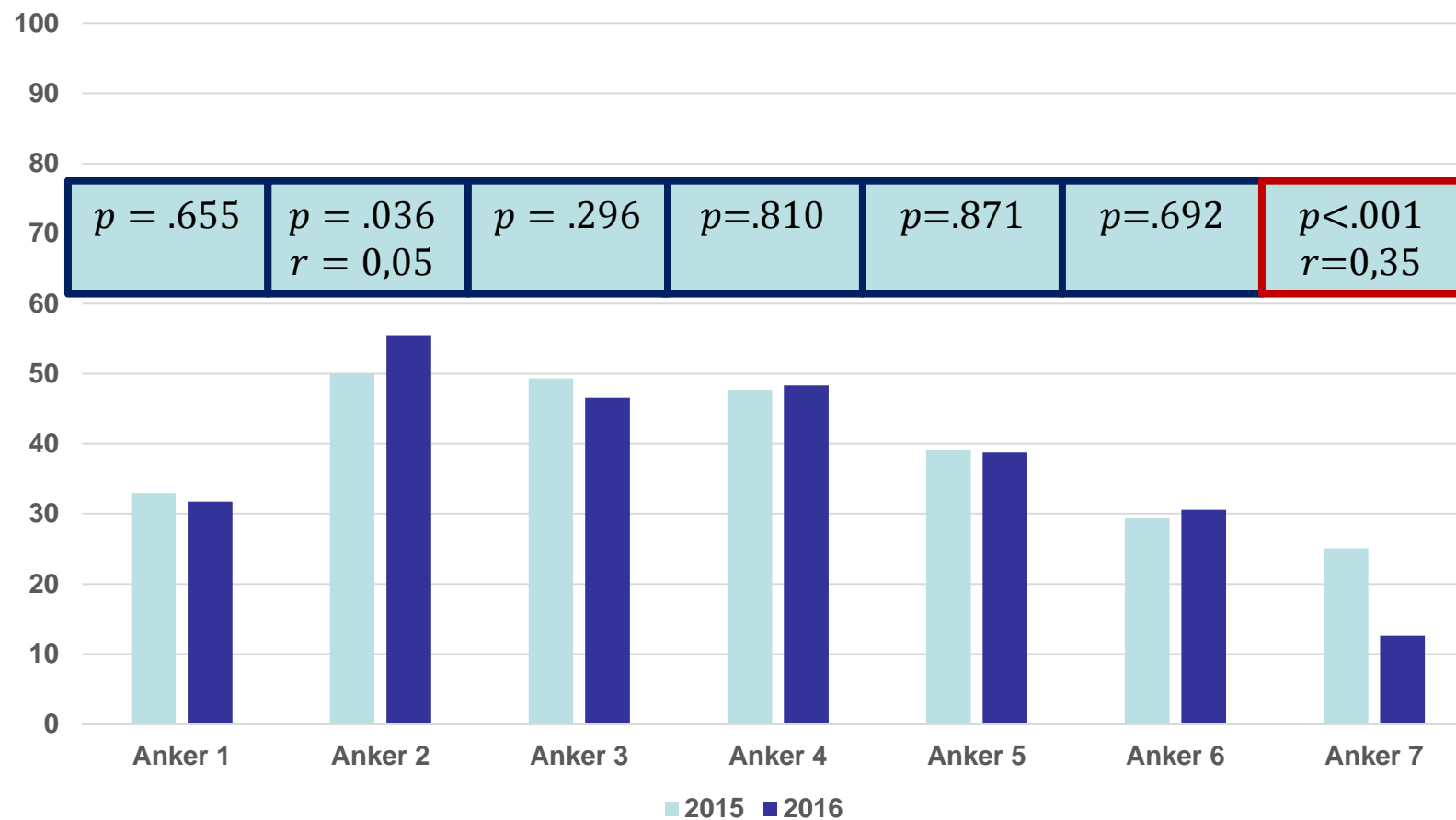
- $\frac{a}{4}\sqrt{3}$
- $\frac{3}{2}a$
- $\sqrt{\frac{3}{2}a^2}$
- $\frac{a}{2}\sqrt{5}$
- Mein Ergebnis ist nicht dabei.
- Ich weiß nicht, wie ich die Aufgabe lösen soll.

■ 2015 ■ 2016



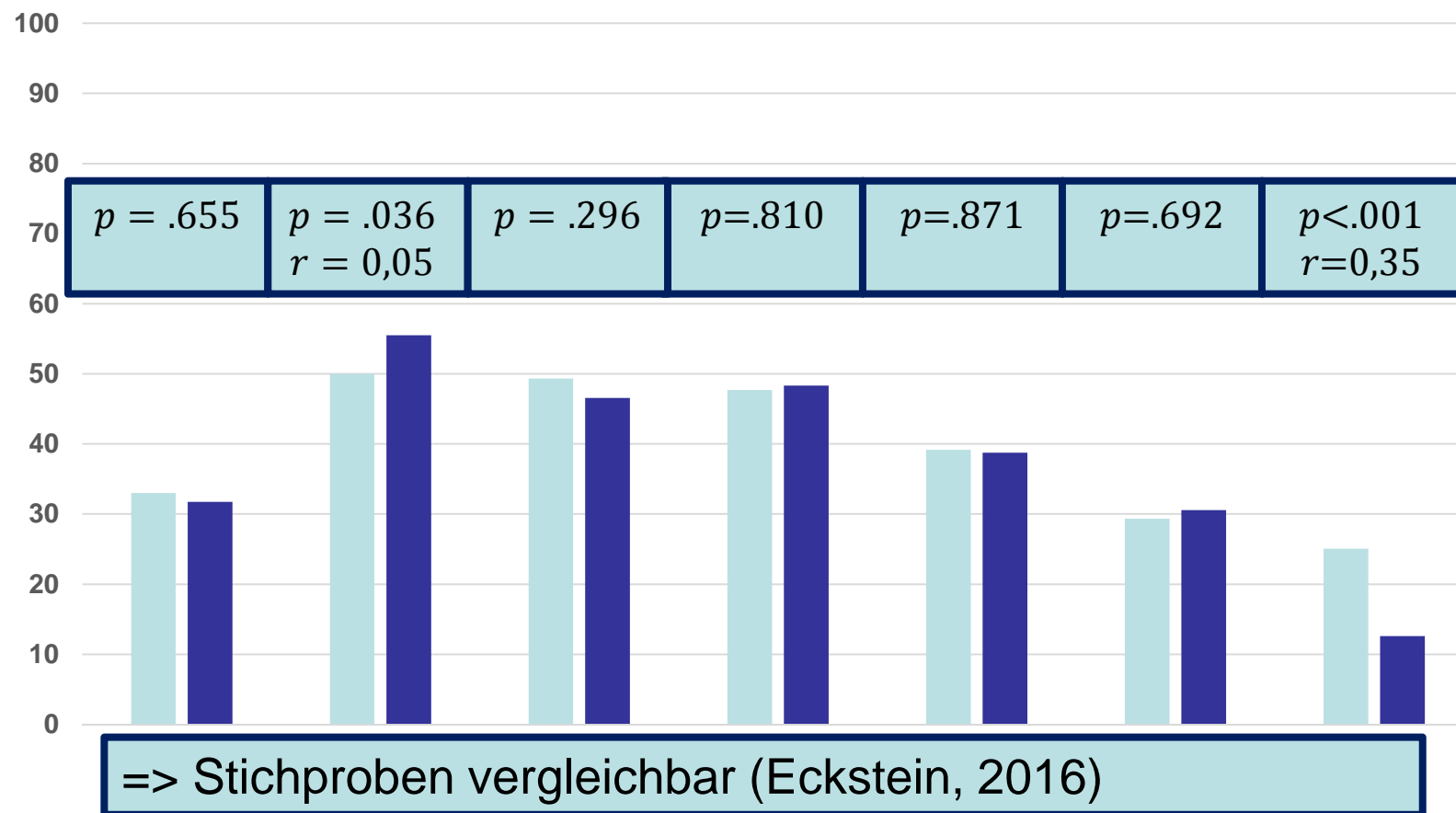
STACK – Chancen und Grenzen

Lösungshäufigkeiten der Ankeritems



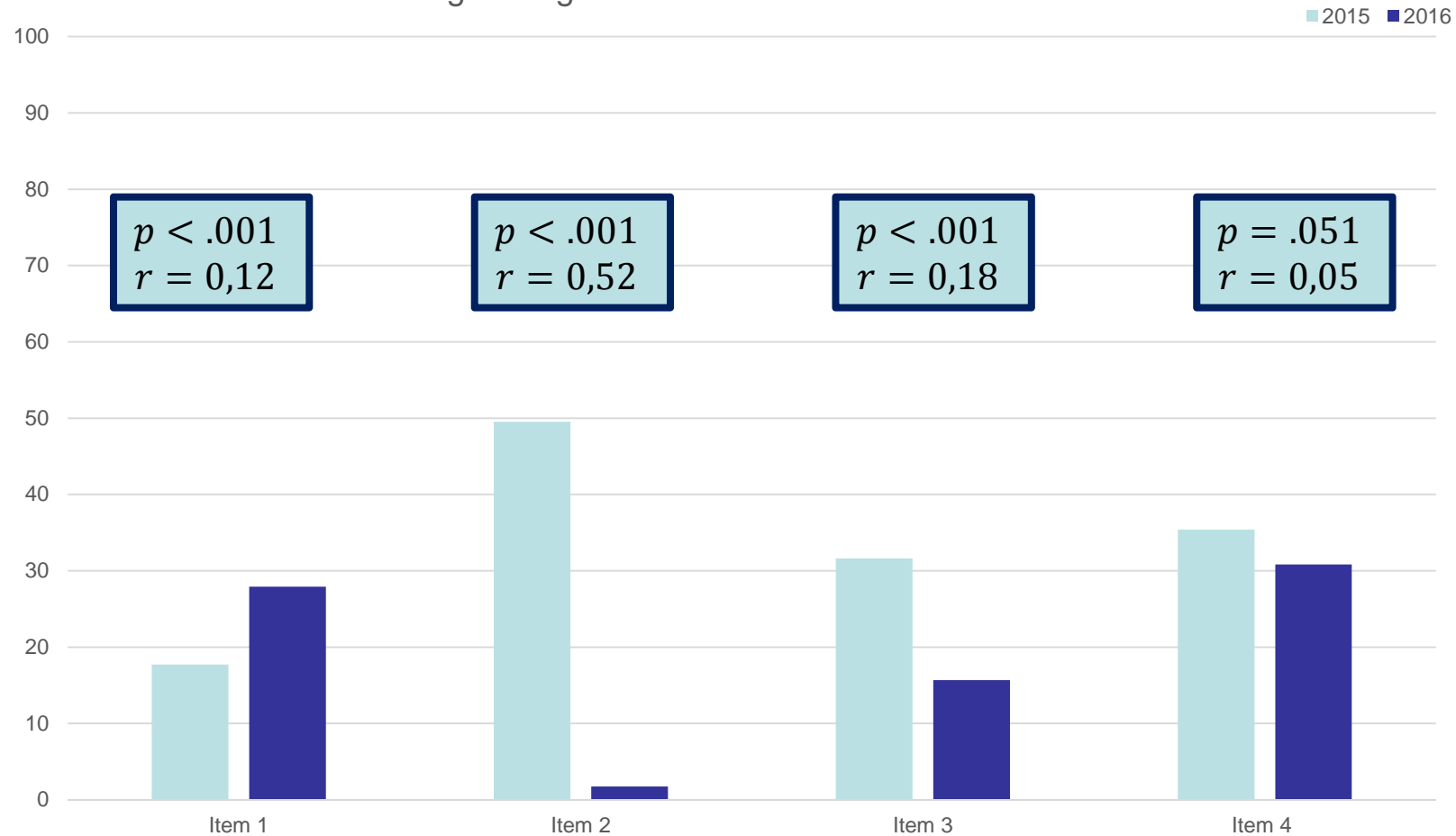
STACK – Chancen und Grenzen

Lösungshäufigkeiten der Ankeritems



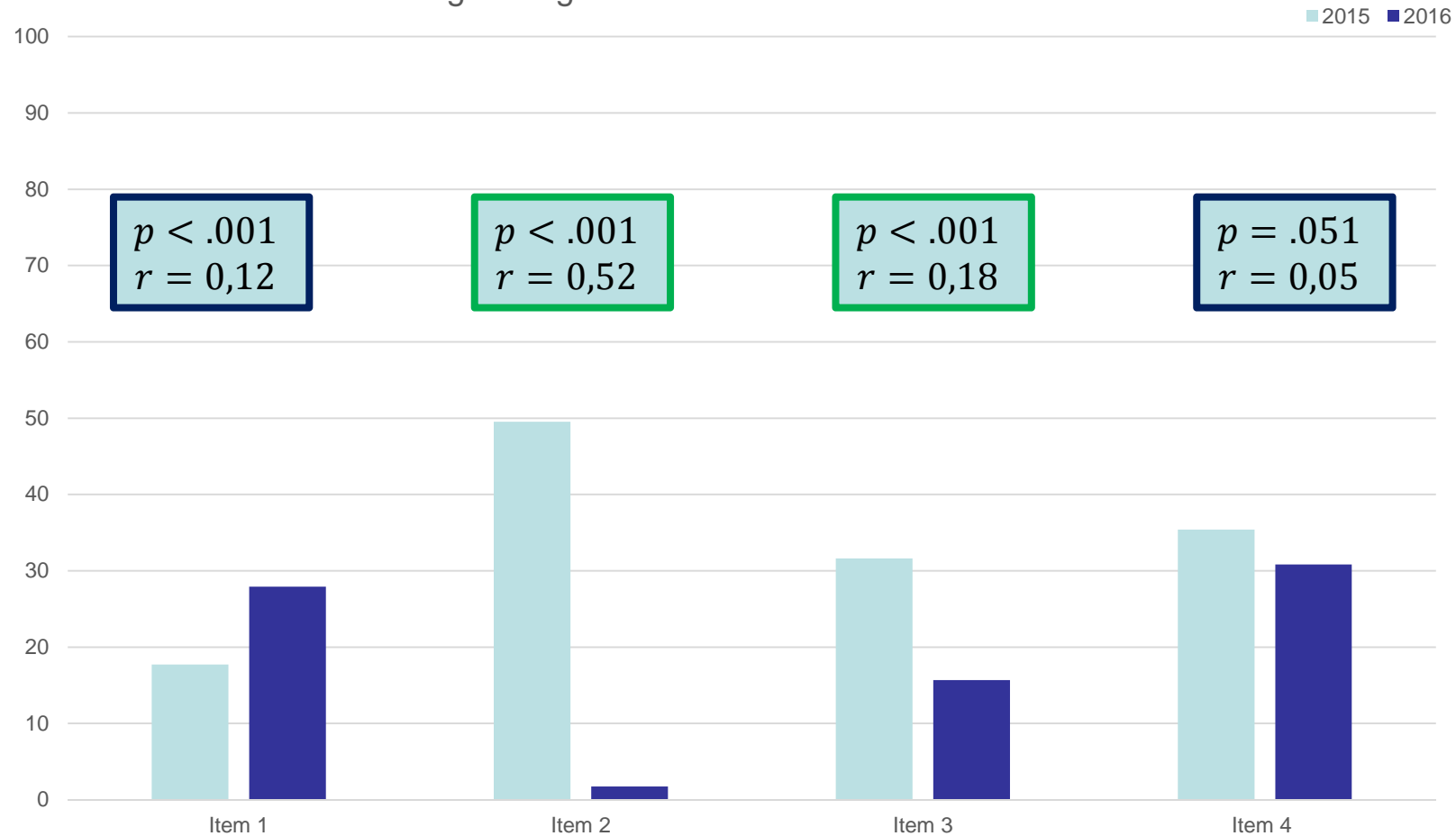
STACK – Chancen und Grenzen

Lösungshäufigkeiten der Items mit Formatwechsel



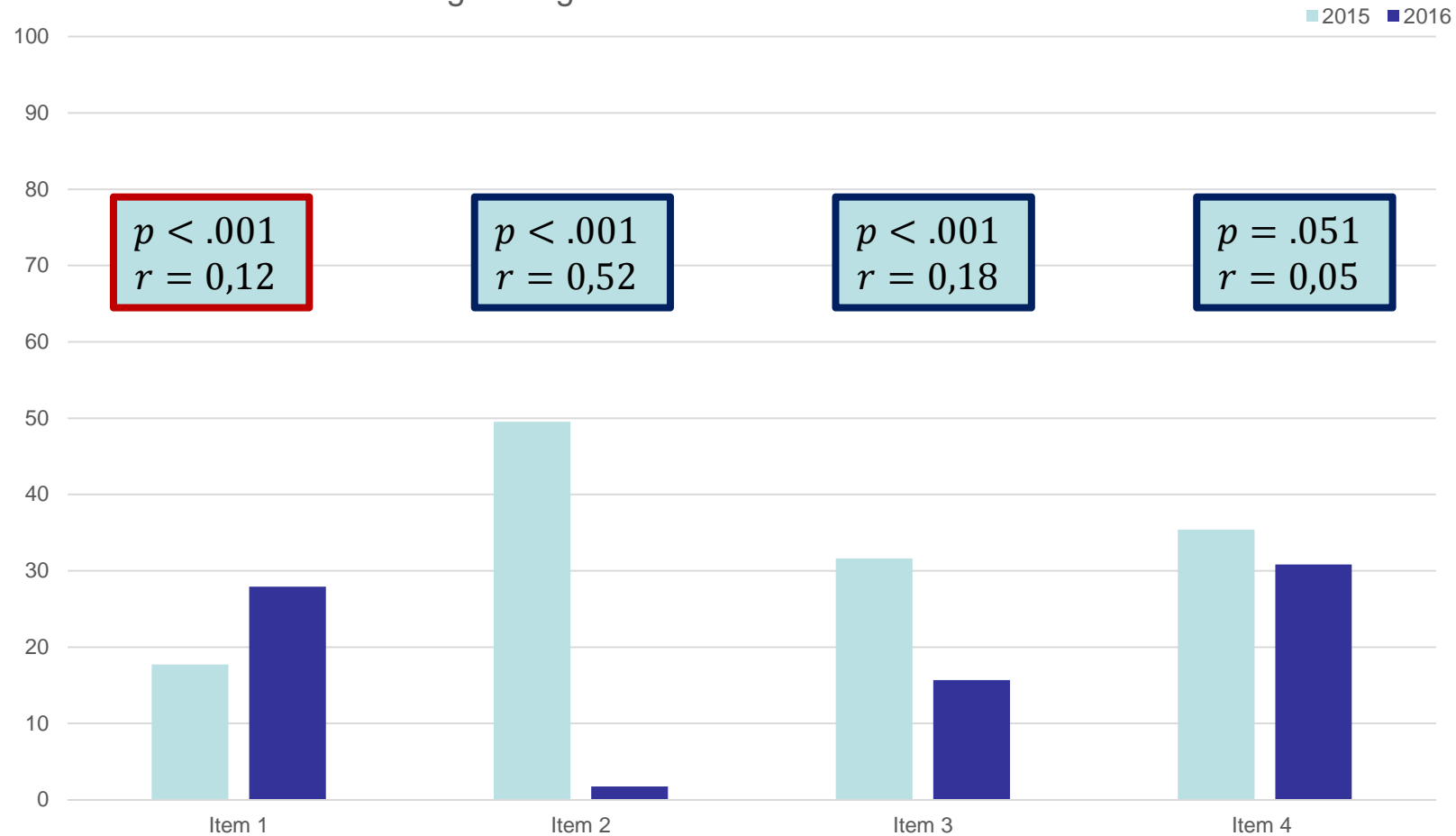
STACK – Chancen und Grenzen

Lösungshäufigkeiten der Items mit Formatwechsel

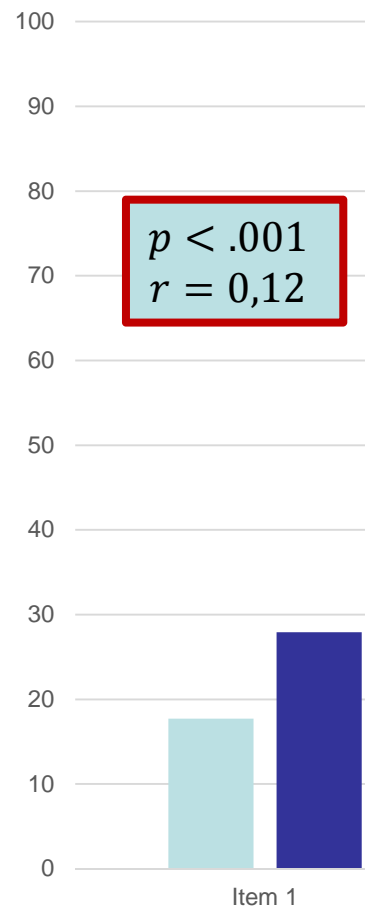


STACK – Chancen und Grenzen

Lösungshäufigkeiten der Items mit Formatwechsel



STACK – Chancen und Grenzen



Welche der folgenden Funktionsgleichungen passt zu dem abgebildeten Funktionsgraphen?

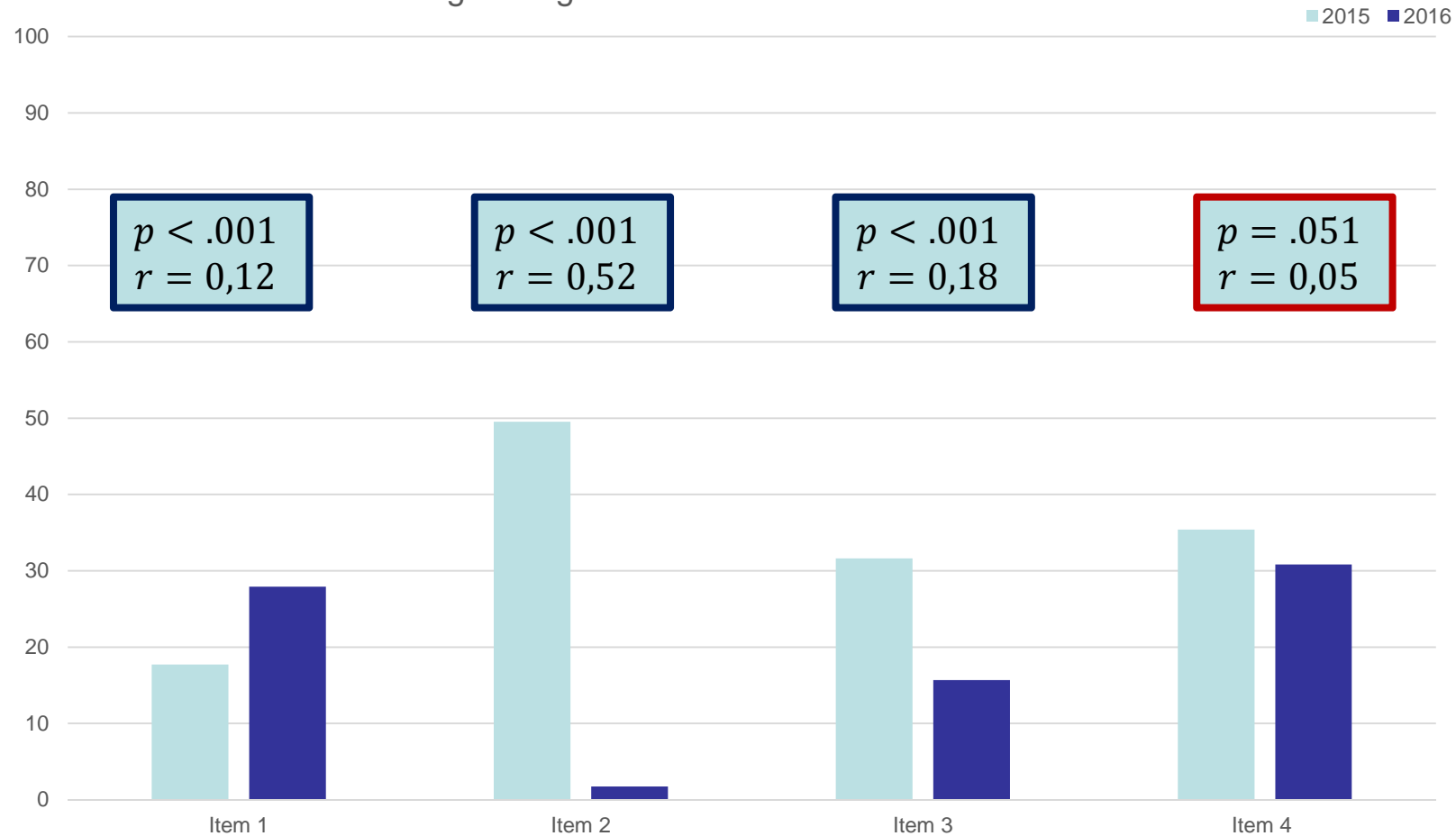
Bitte wählen Sie eine oder mehrere der folgenden Antworten:

Antwort:

- $f(x) = (x - 2)^2 + 3$
- $f(x) = (x - 3)^2 + 2$
- $f(x) = (x + 2)^2 + 3$
- $f(x) = (x + 3)^2 + 2$
- $f(x) = (x + 3)(x + 2)$
- $f(x) = (x - 3)(x + 2)$
- $f(x) = x^2 - 6x + 11$
- $f(x) = x^2 + 6x + 11$
- $f(x) = x^2 + 3x + 2$
- $f(x) = x^2 - 3x + 2$
- Ich weiß nicht, wie ich die Aufgabe lösen soll.

STACK – Chancen und Grenzen

Lösungshäufigkeiten der Items mit Formatwechsel



STACK – Chancen und Grenzen

Lösungshäufigkeiten der Items mit Formatwechsel

100

■ 2015 ■ 2016

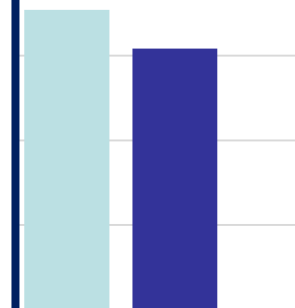
Bestimmen Sie die Ableitungsfunktion der Funktion $f(x) = x^4 \cdot e^x$

Bitte wählen Sie eine der folgenden Antworten:

Antwort:

- $f'(x) = x^4 \cdot e^x$
- $f'(x) = 4x^3 e^x$
- $f'(x) = 4x^3 e^x + x^4$
- $f'(x) = e^x(4x^3 + x^4)$
- $f'(x) = x^4 e^x + 4x^3$
- Mein Ergebnis ist nicht dabei.
- Ich weiß nicht, wie ich die Aufgabe lösen soll.

$p = .051$
 $r = 0,05$



Item 4

STACK – Chancen und Grenzen

Validität

- Keine Ratewahrscheinlichkeit wie beim MC (z.B. Kamps & van Lint, 1975)
- Mehr leere Antworten als über MC mit entsprechender Option (Kallweit, 2016)

Hypothese:

- Aufgaben mit geringer Komplexität sind im offenen Format empirisch schwieriger als im geschlossenen Format, **wenn durch das geschlossene Format keine zusätzlichen Anforderungen notwendig werden.**
- Besonders einschrittige Aufgaben können durch ein Identifizieren statt einem Realisieren der richtigen Antwort gelöst werden (Feld-Caesar, 2017)

STACK – Chancen und Grenzen



Chancen:

- Validere Prüfung durch mehr Aufgaben im offenen Format
- Neue Aufgabenformate
- Automatische Auswertung und Datenerhebung
- Abgleich von typischen Fehlern, Fehlerclustern im Hintergrund
- Individuelles Feedback

STACK – Chancen und Grenzen

Gegeben ist die Funktionsgleichung einer Funktion f durch:

$$f(x) = 2x^2 - 7$$

Bestimmen Sie die Funktionsgleichung der Tangente t an den Graphen der Funktion f im Punkt $P(-2 | 1)$.

N = 303 (WS 2018/2019)

Lösungshäufigkeit: 50,8%

Typische Fehlerphänomene:

- $4x+9$ (7,4% aller Fehler)
- $-8x+17$ (2,7% aller Fehler)

STACK – Chancen und Grenzen

Gegeben ist die Funktionsgleichung einer Funktion f durch:

$$f(x) = 2x^2 - 7$$

Bestimmen Sie die Funktionsgleichung der Tangente t an den Graphen der Funktion f im Punkt $P(-2 | 1)$.

$N = 303$

Lösungshäufigkeit: 50,8%

Typische F

Antwortklasse mit nur korrekter Steigung: $t(x) = -8x + b$
mit $b \neq -15$
Anteil an allen Fehlern: 20,1%

- $4x+9$ (7,4% aller Fehler)
- $-8x+17$ (2,7% aller Fehler)

STACK – Chancen und Grenzen

Chancen:

- Validere Prüfung durch mehr Aufgaben im offenen Format
- Neue Aufgabenformate
- Automatische Auswertung und Datenerhebung
- Abgleich von typischen Fehlern, Fehlerclustern im Hintergrund
- Individuelles Feedback

Grenzen:

- Zu viele untypische Fehler

STACK – Chancen und Grenzen

Vereinfachen Sie den folgenden Term so weit wie möglich und fassen Sie die Variablen zusammen:

$$\frac{x^{k-n}}{y^{2n}} : \frac{x^{2k-n}}{y^{n-1}} \quad (x, y \neq 0)$$

N = 306 (WS 2018/2019)

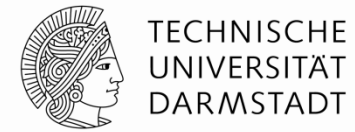
Lösungshäufigkeit: 23,9%

Typische Fehlerphänomen: $\frac{x^{k-n}y^{n-1}}{x^{2k-n}y^{2n}}$ (8,1% aller Fehler)

Anzahl verschiedener Fehler: 96

Anzahl Fehlerunikate: 77

STACK – Chancen und Grenzen



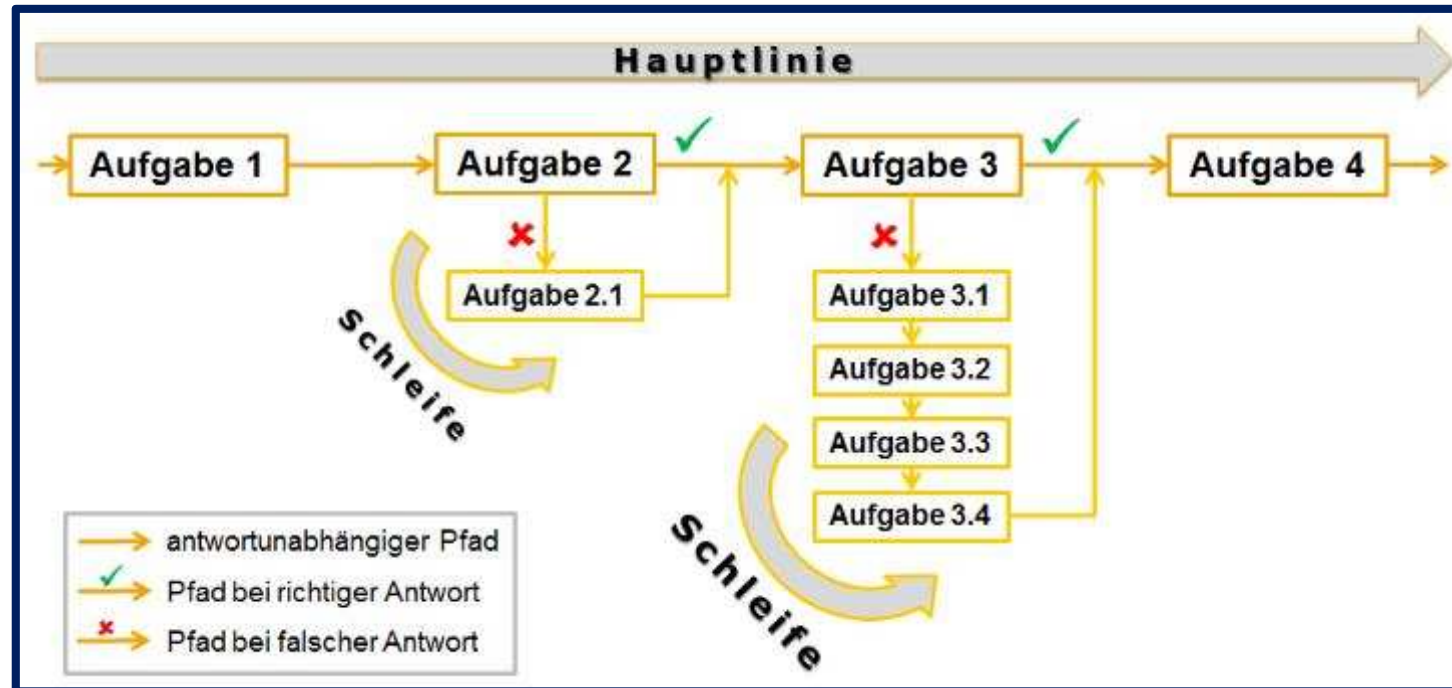
Chancen:

- Validere Prüfung durch mehr Aufgaben im offenen Format
- Neue Aufgabenformate
- Automatische Auswertung und Datenerhebung
- Abgleich von typischen Fehlern, Fehlerclustern im Hintergrund
- Individuelles Feedback

Grenzen:

- Zu viele untypische Fehler
- Prüfen verknüpfter Lerninhalte – im Sinne intelligenten Wissens (Weinert, 2000) notwendig (Feldt-Caesar, 2017)

STACK – Chancen und Grenzen



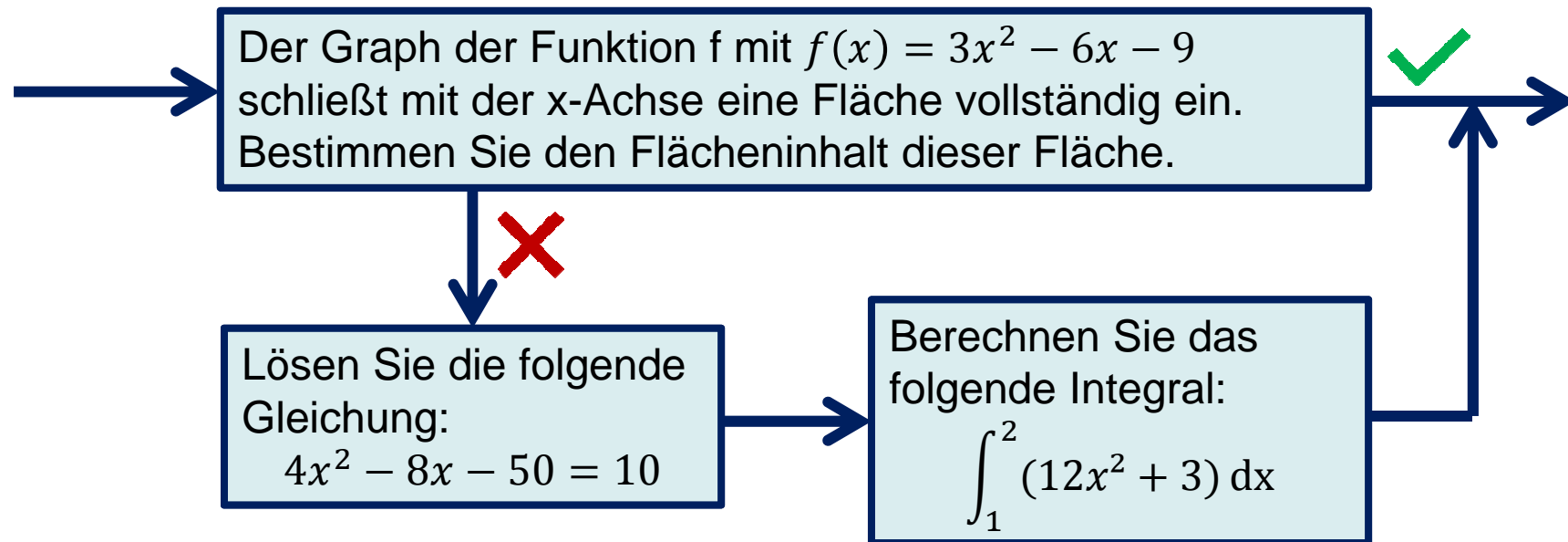
(Feldt-Caesar, 2017)

STACK – Chancen und Grenzen

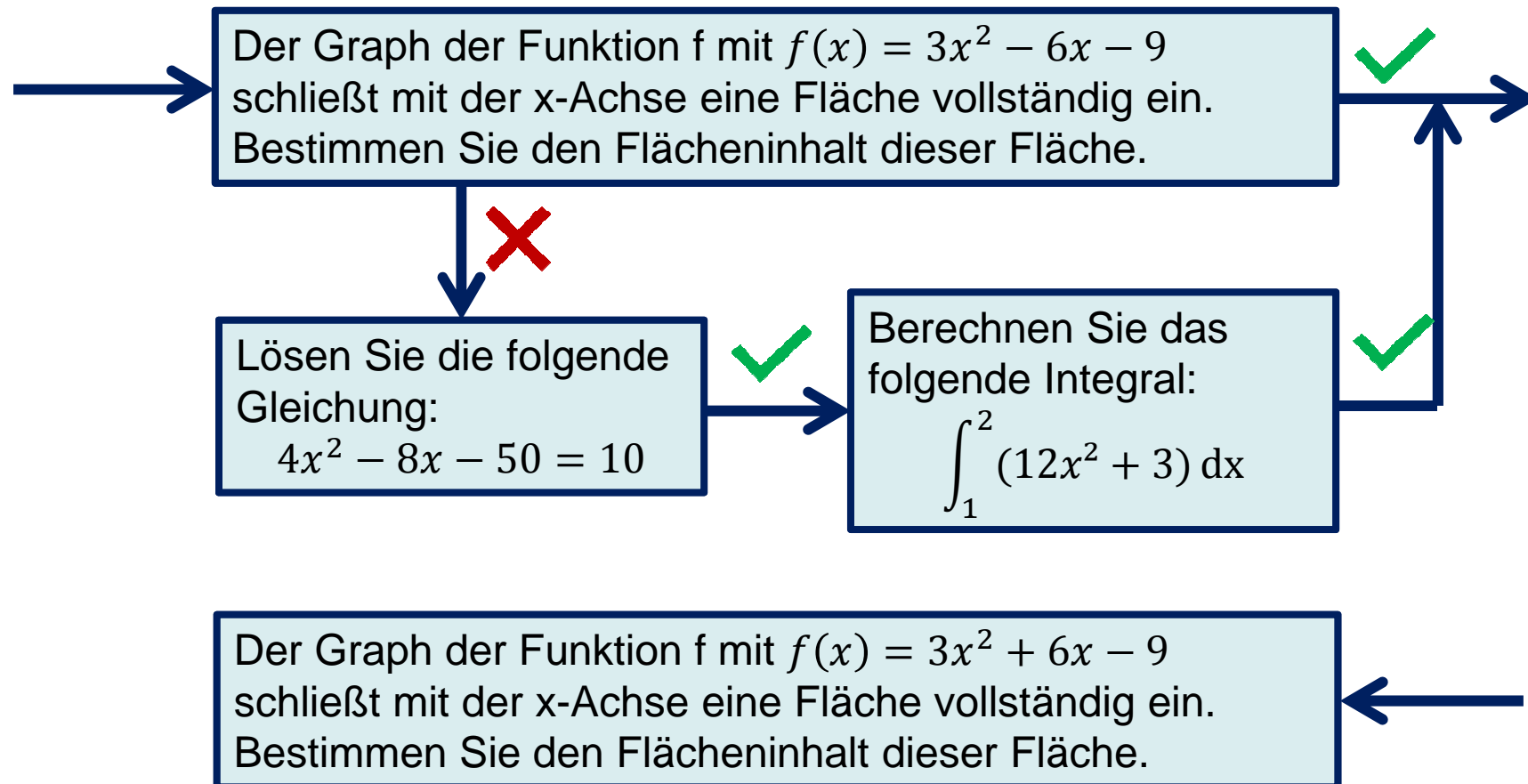
Der Graph der Funktion f mit $f(x) = 3x^2 - 6x - 9$ schließt mit der x -Achse eine Fläche vollständig ein. Bestimmen Sie den Flächeninhalt dieser Fläche.



STACK – Chancen und Grenzen



STACK – Chancen und Grenzen



STACK – Chancen und Grenzen



-
- Neu entwickeltes moodle-Plugin im Rahmen eines Bachelor-Praktikums Informatik, derzeit für moodle 3.1
 - Ermöglicht neue Testaktivität DDTA-Quiz
 - Verwendet Fragen der klassischen Testaktivität
 - Adaptives Feedback
 - Einsatz im Ein- und Ausgangstest in Darmstadt im WS18/19

Elementarisierendes Testen

- + Frage hinzufügen
- + aus der Fragesammlung
- + **Block hinzufügen**

16. Schleife A8(p)
16. Vtest_18_S8.1
17. Vtest_18_S8.2

18. Vtest_18_A9(p)

19. VTest_18_A10(p)

20. Vtest_18_A11(p)

21. Vtest_18_A12(p)

22. Vtest_18_A13(p)

23. Schleife A13(p)
23. Vtest_18_S13.1
24. Vtest_18_S13.2
25. Vtest_18_S13.3

26. Schleife S13.3
26. Vtest_18_S13.3.1

27. VTest_18_S13.4

Elementarisierendes Testen

Bedingungen

Ein Teilnehmer muss der Bedingungen erfüllen,
um diesen Block zu sehen.

Punkte aus sind =

Punkte aus sind =

Punkte aus sind =

[Bedingung hinzufügen](#)

Elementarisierendes Testen

**Dieses Feedback ersetzt das Feedback
der folgenden Fragen**

A

B

C

D

[Frage hinzufügen](#)

Elementarisierendes Testen

Bedingungen

Ein Teilnehmer muss der Bedingungen erfüllen,
um diesen Block zu sehen.

Punkte aus sind =

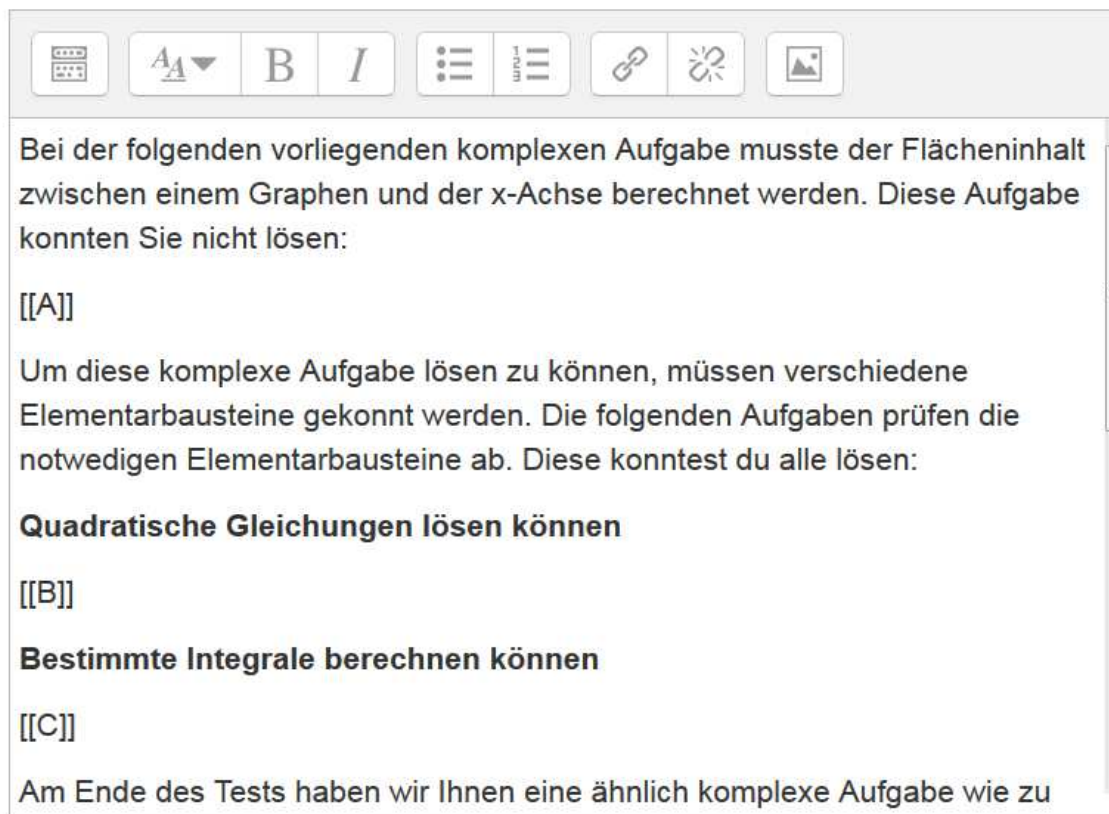
Punkte aus sind =

Punkte aus sind =

[Bedingung hinzufügen](#)

Elementarisierendes Testen

Feedbacktext



The screenshot shows a text editor interface with a toolbar at the top containing icons for text alignment, bold, italic, list, link, unlink, and image. The main text area contains the following content:

Bei der folgenden vorliegenden komplexen Aufgabe musste der Flächeninhalt zwischen einem Graphen und der x-Achse berechnet werden. Diese Aufgabe konnten Sie nicht lösen:

[[A]]

Um diese komplexe Aufgabe lösen zu können, müssen verschiedene Elementarbausteine gekonnt werden. Die folgenden Aufgaben prüfen die notwendigen Elementarbausteine ab. Diese konntest du alle lösen:

Quadratische Gleichungen lösen können

[[B]]

Bestimmte Integrale berechnen können

[[C]]

Am Ende des Tests haben wir Ihnen eine ähnlich komplexe Aufgabe wie zu

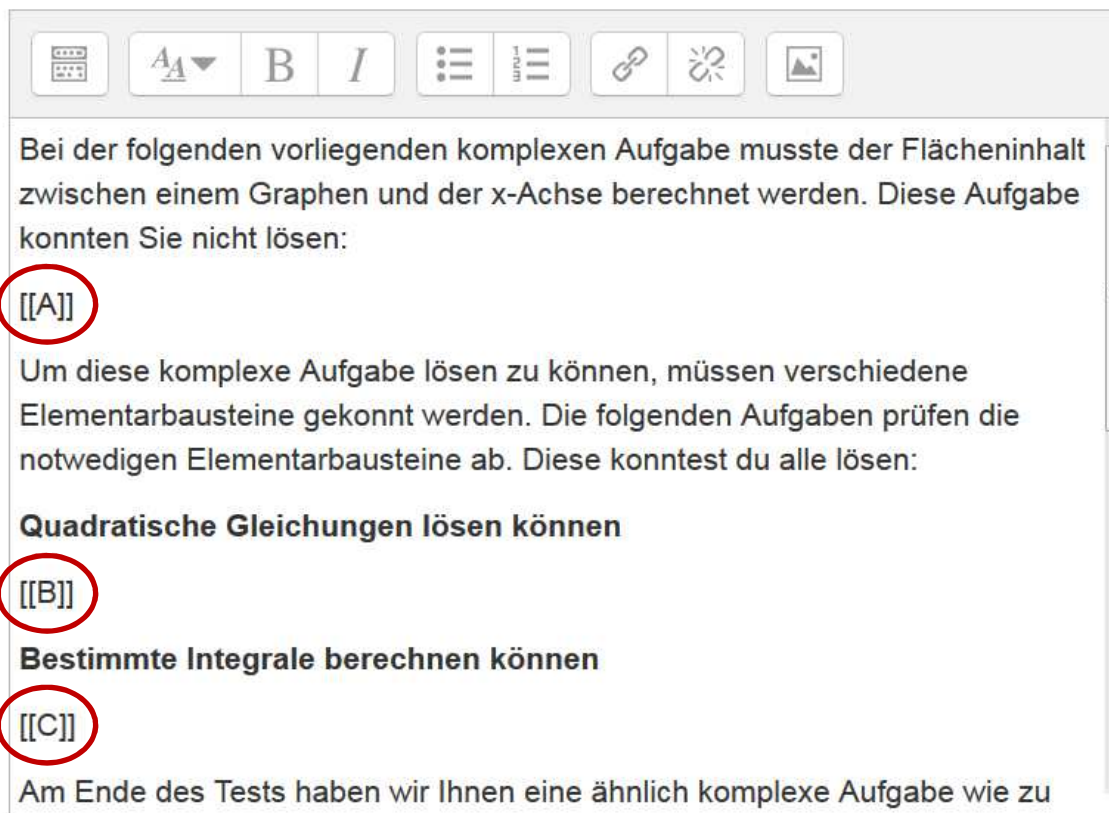
Elementarisierendes Testen

**Dieses Feedback ersetzt das Feedback
der folgenden Fragen**

- A** 2. Vortest_17_A2 x
- B** 3. Vortest_17_S2.1 x
- C** 4. Vortest_17_S2.2 x
- D** 8. VTest_17_P2 x

[Frage hinzufügen](#)

Feedbacktext



The screenshot shows a rich text editor interface. At the top, there is a toolbar with icons for text alignment, bold, italic, bulleted list, numbered list, link, unlink, and image. Below the toolbar, the text reads: "Bei der folgenden vorliegenden komplexen Aufgabe musste der Flächeninhalt zwischen einem Graphen und der x-Achse berechnet werden. Diese Aufgabe konnten Sie nicht lösen:" followed by a red circle around "[[A]]". The next paragraph says: "Um diese komplexe Aufgabe lösen zu können, müssen verschiedene Elementarbausteine gekonnt werden. Die folgenden Aufgaben prüfen die notwendigen Elementarbausteine ab. Diese konntest du alle lösen:" followed by the bold heading "Quadratische Gleichungen lösen können", a red circle around "[[B]]", the bold heading "Bestimmte Integrale berechnen können", and a red circle around "[[C]]". The final sentence is: "Am Ende des Tests haben wir Ihnen eine ähnlich komplexe Aufgabe wie zu".

STACK – Chancen und Grenzen

Fehleraufklärungsquote (FQ)

$$FQ = \frac{N_{AF}}{N_F}$$

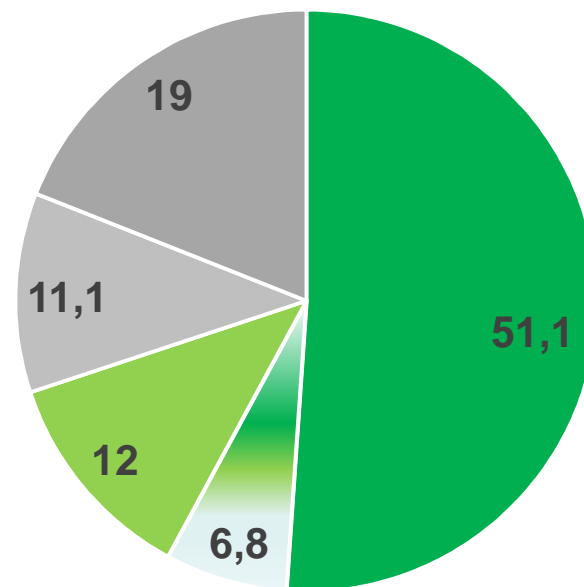
N_F := Anzahl aller falschen Antworten zu einer Aufgabe

N_{AF} := Anzahl falscher Antworten, die zumindest teilweise aufgeklärt werden konnten

(Schaub, 2018; modifiziert nach Feldt-Caesar, 2017)

Elementarisierendes Testen

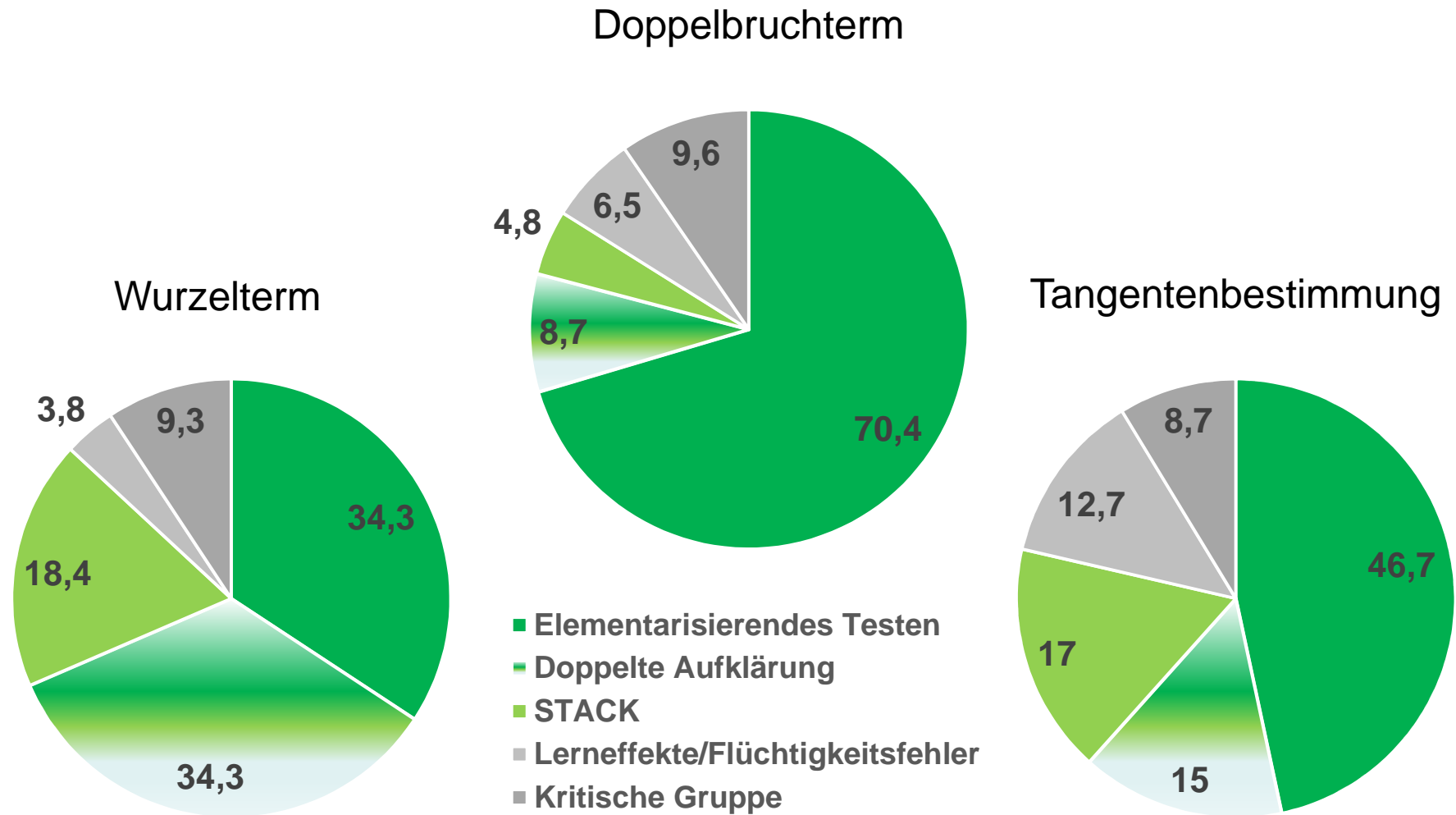
Aufklärung der Fehler im Beispiel “Flächeninhalt” in %



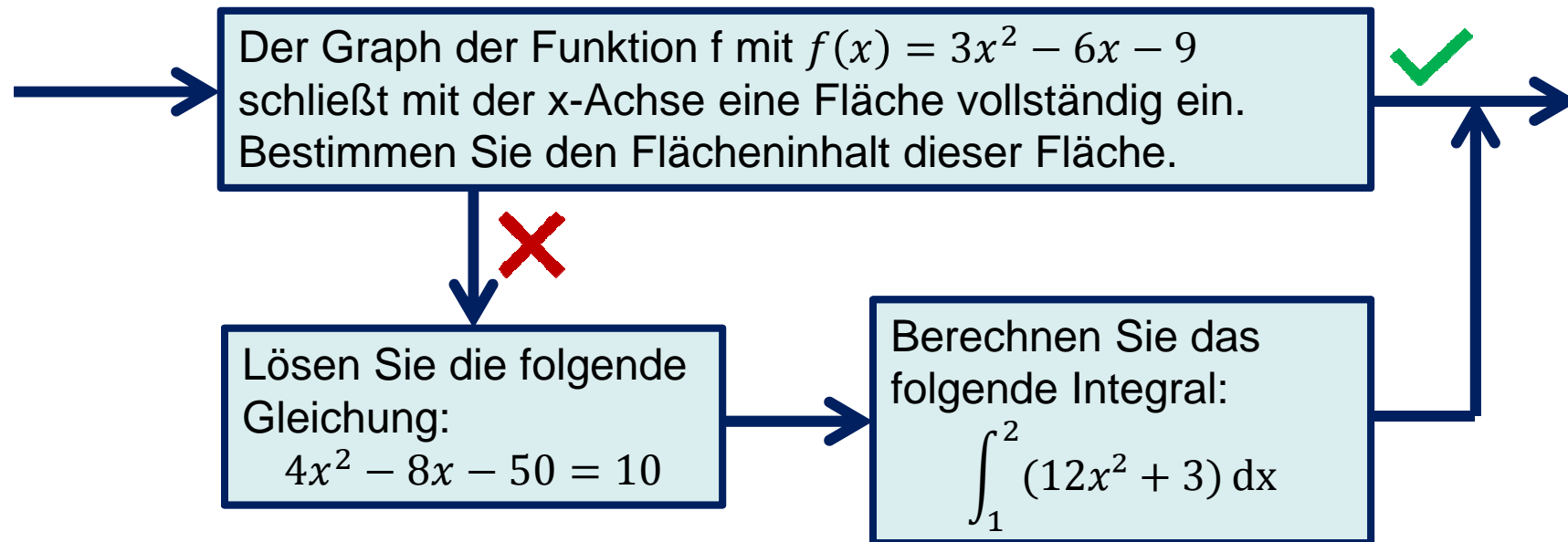
■ Elementarisierendes Testen
■ STACK
■ Kritische Gruppe

■ Doppelte Aufklärung
■ Lerneffekte/Flüchtigkeitsfehler

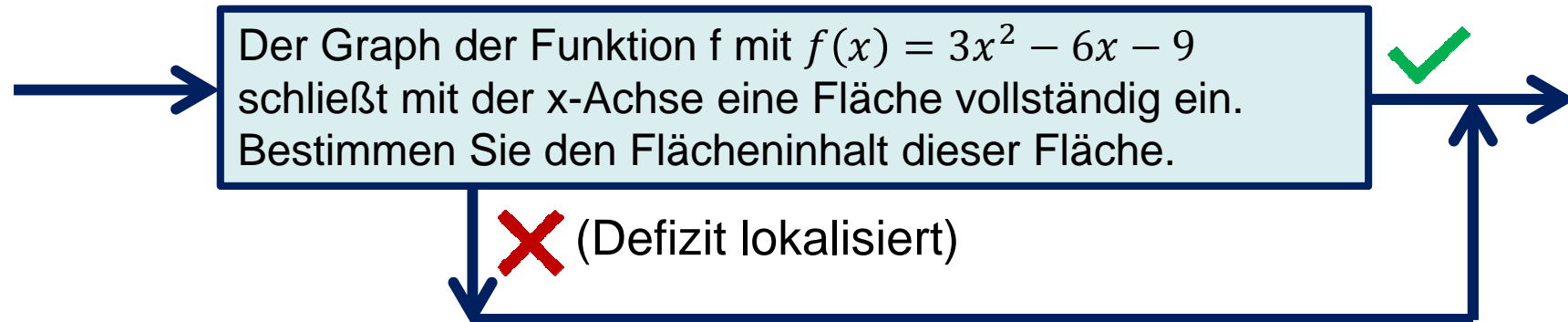
Elementarisierendes Testen



STACK – Chancen und Grenzen



STACK – Chancen und Grenzen



- Fertigstellung Plugin für weitere Moodle-Versionen inkl. neuer Funktionen
- Verknüpfung beider Ansätze dauerhaft implementieren, da gewinnbringende Diagnoseinformationen
- Zusammenfassende Analyse mehrerer Jahrgänge zur Bestimmung weiterer typischer Fehler
- Analyse zur Konsistenz von Fehlern in den elementarisierenden Schleifen

Vielen Dank für die Aufmerksamkeit!



Literatur



- Bausch, I.; Biehler, R.; Bruder, R.; Fischer, P. R.; Hochmuth, R.; Koepf W. et al. (Hg.) (2014): Mathematische Vor- und Brückenkurse. Konzepte, Probleme und Perspektiven. Wiesbaden: Springer Fachmedien Wiesbaden (Konzepte und Studien zur Hochschuldidaktik und Lehrerbildung Mathematik).
- Beiträge zum Mathematikunterricht 2018. Vorträge zur Mathematikdidaktik und zur Schnittstelle Mathematik/Mathematikdidaktik auf der gemeinsamen Jahrestagung GDM und DMV 2018 (52. Jahrestagung der Gesellschaft für Didaktik der Mathematik) (2018). 1. Erste Auflage. Münster: WTM-Verlag.
- Biehler, R.; Bruder, R.; Hochmuth, R.; Koepf W. (2014): Einleitung. In: I. Bausch, R. Biehler, R. Bruder, P. R. Fischer, R. Hochmuth, Koepf W. et al. (Hg.): Mathematische Vor- und Brückenkurse. Konzepte, Probleme und Perspektiven. Wiesbaden: Springer Fachmedien Wiesbaden (Konzepte und Studien zur Hochschuldidaktik und Lehrerbildung Mathematik), S. 1–6.
- Biehler, R.; Bruder, R.; Hochmuth, R.; Koepf, W.; Bausch, I.; Fischer, P. R.; Wassong T. (2014): VEMINT - Interaktives Lernmaterial für mathematische Vor- und Brückenkurse. In: I. Bausch, R. Biehler, R. Bruder, P. R. Fischer, R. Hochmuth, Koepf W. et al. (Hg.): Mathematische Vor- und Brückenkurse. Konzepte, Probleme und Perspektiven. Wiesbaden: Springer Fachmedien Wiesbaden (Konzepte und Studien zur Hochschuldidaktik und Lehrerbildung Mathematik), S. 261–276.
- Biehler, R.; Bruder, R.; Hochmuth, R.; Koepf, W.; Bausch, I.; Fischer, P. R.; Wassong T. (2014): VEMINT - Interaktives Lernmaterial für mathematische Vor- und Brückenkurse. In: I. Bausch, R. Biehler, R. Bruder, P. R. Fischer, R. Hochmuth, Koepf W. et al. (Hg.): Mathematische Vor- und Brückenkurse. Konzepte, Probleme und Perspektiven. Wiesbaden: Springer Fachmedien Wiesbaden (Konzepte und Studien zur Hochschuldidaktik und Lehrerbildung Mathematik), S. 261–276.
- Eckstein, Peter P. (2016): Angewandte Statistik mit SPSS. Praktische Einführung für Wirtschaftswissenschaftler. 8., überarb. u. erw. Aufl. 2016. Wiesbaden: Springer Fachmedien Wiesbaden GmbH; Springer Gabler (Lehrbuch).
- Feldt-Caesar, Nora (2017): Konzeptualisierung und Diagnose von Mathematischem Grundwissen und Grundkönnen. Eine Theoretische Betrachtung und Exemplarische Konkretisierung am Ende Der Sekundarstufe II: Spektrum Akademischer Verlag GmbH.
- Hasselhorn, Marcus (2014): Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen. Bielefeld: Wbv, Bertelsmann (Bildung in Deutschland, 2014).
- Kallweit, Michael (2016): Der Computer als Tutor - technikbasierte Diagnostik mit Freitextaufgaben. In: Beiträge zum Mathematikunterricht 2016. Vorträge auf der 50. Tagung für Didaktik der Mathematik vom 07.03.2016 bis 11.03.2016 in Heidelberg. Münster: WTM - Verlag für wissenschaftliche Texte und Medien, S. 1153–1156.
- Kamps, H. J. L.; van Lint, J. H. (1975): A comparison of a classical calculus test with a similar multiple choice test. In: *Educ Stud Math* 6 (3), S. 259–271. DOI: 10.1007/BF01793611.

Literatur



Katz, I. R.; Friedman, D. E.; Bennett, R. E.; Berger, A. E. (1996): Differences in Strategies Used to Solve Stem-Equivalent Constructed-Response and Multiple-Choice SAT[®]-Mathematics Items. In: *ETS Research Report Series 1996 (2)*, i-20. DOI: 10.1002/j.2333-8504.1996.tb01698.x.

Mathematik-Kommission Übergang Schule-Hochschule (20.04.2017): Zur aktuellen Diskussion über die Qualität des Mathematikunterrichts. Online verfügbar unter <http://www.mathematik-schule-hochschule.de/images/Stellungnahmen/pdf/Stellungnahme-DMVGDMMNU-2017.pdf>, zuletzt geprüft am 07.12.2018.

Krüger-Basener, M.; Rabe, D. (2014): Mathe0 - der Einführungskurs für *alle* Erstsemester einer technischen Lehreinheit. In: I. Bausch, R. Biehler, R. Bruder, P. R. Fischer, R. Hochmuth, Koepf W. et al. (Hg.): *Mathematische Vor- und Brückenkurse. Konzepte, Probleme und Perspektiven*. Wiesbaden: Springer Fachmedien Wiesbaden (Konzepte und Studien zur Hochschuldidaktik und Lehrerbildung Mathematik), S. 309–323.

Krüger-Basener, M.; Rabe, D. (2014): Mathe0 - der Einführungskurs für *alle* Erstsemester einer technischen Lehreinheit. In: I. Bausch, R. Biehler, R. Bruder, P. R. Fischer, R. Hochmuth, Koepf W. et al. (Hg.): *Mathematische Vor- und Brückenkurse. Konzepte, Probleme und Perspektiven*. Wiesbaden: Springer Fachmedien Wiesbaden (Konzepte und Studien zur Hochschuldidaktik und Lehrerbildung Mathematik), S. 309–323.

Nitsch, Renate (2015): *Diagnose von Lernschwierigkeiten im Bereich funktionaler Zusammenhänge. Eine Studie zu typischen Fehlermustern bei Darstellungswechseln*. Wiesbaden: Springer Fachmedien Wiesbaden.

Pädagogische Nachrichten Rheinland-Pfalz (2000).

Sangwin, Christopher J.; Jones, Ian (2017): Asymmetry in student achievement on multiple-choice and constructed-response items in reversible mathematics processes. In: *Educ Stud Math* 94 (2), S. 205–222. DOI: 10.1007/s10649-016-9725-4.

Schaub, Marcel (2018): Einsatz des Elementarisierenden Testens im Ein- und Ausgangstest des online-Vorkurses VEMINT. In: *Beiträge zum Mathematikunterricht 2018. Vorträge zur Mathematikdidaktik und zur Schnittstelle Mathematik/Mathematikdidaktik auf der gemeinsamen Jahrestagung GDM und DMV 2018 (52. Jahrestagung der Gesellschaft für Didaktik der Mathematik)*. 1. Erste Auflage. Münster: WTM-Verlag, S. 1567–1570.

Literatur



Schiemann, Stephanie (2013): Mathe = Mathe? Mathematik in den 16 Bundesländern. In: *Mitteilungen der Deutschen Mathematiker-Vereinigung* 21 (4). DOI: 10.1515/dmvm-2013-0087.

Tagung für Didaktik der Mathematik; Jahrestagung der Gesellschaft für Didaktik der Mathematik (2016): Beiträge zum Mathematikunterricht 2016. Vorträge auf der 50. Tagung für Didaktik der Mathematik vom 07.03.2016 bis 11.03.2016 in Heidelberg. Münster: WTM - Verlag für wissenschaftliche Texte und Medien.

Weinert, F. E. (2000): Lehren und Lernen für die Zukunft - Ansprüche an das Lernen in der Schule. In: *Pädagogische Nachrichten Rheinland-Pfalz*, Bd. 2, Sonderseiten 1-16. Online verfügbar unter <http://www2.ibw.uni-heidelberg.de/~gerstner/WeinertLehren&Lernen.pdf>, zuletzt geprüft am 31.01.2018.

Backup - Elementarisierendes Testen



Fehleraufklärungsquote (alte Definition)

Gruppe 2 (Lerneffekte)

Kritische Gruppe

(Feldt-Caesar, 2017)