



# Do the Tenets of Answer Changing Research Hold for an Innovative Assessment?



Katherine E. Castellano, Jamie N. Mikeska, Jung Aa Moon, Steven Holtzman, Jie Gao, & Yang Jiang, ETS

## Background

- There is a common belief that examinees should stick to their original answers.
- But a large body of research (e.g., Mathews, 1929; Waddell & Blankenship, 1994; Al-Hamly & Coombe, 2004) indicates that examinees make more beneficial (wrong-to-right; WR) than detrimental changes (right-to-wrong; RW) and, thus, more often gain than lose points by changing their answer.
- There is no or little research on answer changing behavior for:
  - The pre-service teacher (PST) test-taker population,
  - Tests assessing integrated constructs of content and practice, or
  - Across different technology-enhanced item types.
- Our study is a first step in filling in these gaps in the answer changing literature. We use a research instrument:
  - Administered online to 822 PSTs,
  - That assesses content knowledge for teaching (CKT) for matter and its interactions,
  - Using several item types, including single-select multiple choice, multiple-select multiple choice, and grid items.

## Research Questions

1. How does PSTs' answer changing behavior relate to (a) item-level and (b) total test scores on the CKT matter test?
2. To what extent does answer changing behavior differ across PSTs' performance levels?

## Sample and Instrument

- **Sample:** 822 PSTs drawn from the Praxis Elementary Science test-taker population
  - Representative of Praxis test-takers in terms of gender, geographic region, race/ethnicity, and Praxis Science performance (by quartiles)
- **Instrument:** Assessed CKT for matter and its interactions—integrated construct of content knowledge and teaching practices
  - Focused on 48 (of 60) items of relevant item types:
    - 24 Single-Select Multiple Choice (SSMC)
    - 17 Multiple-Select Multiple Choice (MSMC)
    - 7 Grid items (Grid)

## Contact

Katherine Furgol Castellano Senior Research Scientist ETS  
 KCastellano@ets.org

Jamie N. Mikeska Senior Research Scientist ETS  
 jmikeska@ets.org

## 1a. How does PSTs' answer changing behavior relate to item-level scores on the CKT matter test?

	All Item Types (48 items)	SSMC (24 items)	MSMC (17 items)	Grid (7 items)
Mean percent of examinees changing a response on an item	19.0	15.9	20.8	25.5
RW	2.7	2.8	2.7	2.5
WR	<b>8.4</b>	<b>9.1</b>	<b>7.7</b>	8.0
WW	5.9	2.6	7.6	<b>13.1</b>
RR	2.0	1.5	2.8	1.9
Item gain-to-loss ratio (WR/RW)	3.1	3.3	2.9	3.2

RW=right-to-wrong; WR = wrong-to-right; WW = wrong-to-wrong; RR = right-to-right; SSMC=single-select multiple-choice items; MSMC=multiple-selection multiple-choice items.

- For the 3 item types, **8-9%** of test-takers, on average, changed responses from wrong to right on each item, resulting in item-level score increases.
- For the 3 item types, about **three times** as many test-takers changed responses from wrong to right over right to wrong on each item.
- For grid items, more examinees changed responses wrong to wrong than any other response pattern change.
  - But grid items were dichotomously scored and a test-taker could fix the response for one row of the grid but not another row so would result in a wrong-to-wrong change.

## 2. To what extent does answer changing behavior differ across PSTs' performance levels?

- Test-takers were classified by quartiles of their total score on all 60 items.
- Higher performers benefited more from answer changing than lower performers.
- Higher performers were 3 to 8 times more likely to make score gains than score losses.
- Lower performers were still more likely to make score gains than losses but not to the same extent as high performers: only 0.5 to 3 times more likely.

## References

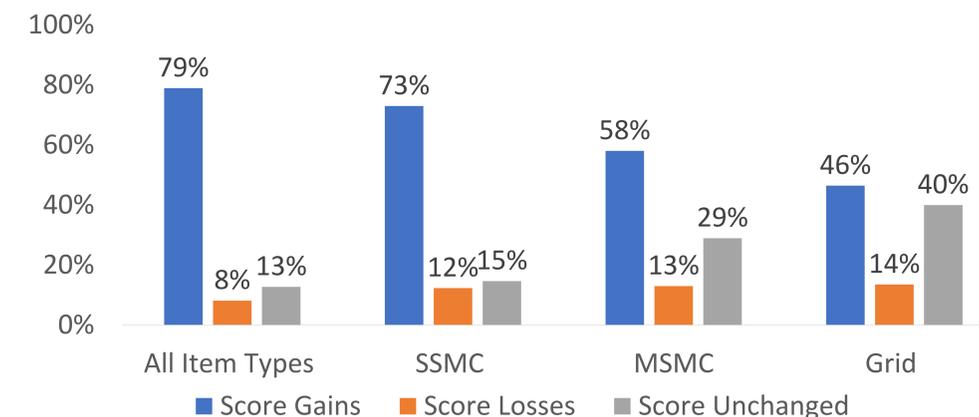
Al-Hamly, M., & Coombe, C. (2005). To change or not to change: Investigating the value of MCQ answer changing for Gulf Arab students. *Language Testing*, 22(4), 509-531.

Mathews, C. O. (1929). Erroneous first impressions on objective tests. *Journal of Educational Psychology*, 20(4), 280.

Waddell, D. L., & Blankenship, J. C. (1994). Answer changing: A meta-analysis of the prevalence and patterns. *The Journal of Continuing Education in Nursing*, 25(4), 155-158.

## 1b. How does PSTs' answer changing behavior relate to total scores on the CKT matter test?

Percentage of Answer Changers with Different Score Changes by Item Type



- For the 3 item types, about 50 to 70% of answer changers gained points.
- Only about 12-14% of answer changers lost score points.
- About 3 to 6 times as many answer changers made gains versus losses.
- For grid items, there is a higher percentage of answer changers with scores unchanged than for other item types but follows from higher prevalence of WW response changes to items (Question 1a).

## Discussion and Implications

- Yes, the tenets of answer changing research hold for PSTs, for a test assessing an integrated construct of content and teaching practice, and for different item types.
- PSTs for Praxis and other similar teacher licensure exams should be encouraged to change responses (after thoughtful consideration).
- Most noticeable distinction across item types is that grid items resulted in more WW changes and more unchanged final test scores but partly due to dichotomous scoring of these items. Other scoring rules can be examined.
- Further research needed on other similar tests and in more high-stakes environments.

## Acknowledgement

This study was supported by a grant from the National Science Foundation (Award No. 1813254). The opinions expressed herein are those of the authors and not the funding agency.

For sample items, visit our [project website](#).