technische universität
dortmund

# Self-testing in higher education
# What predicts the use of self-testing and who achieves their previously set goals?

Jakob Schwerter[1], Fani Lauermann[1,2], Taiga Brahm[3] & Kou Murayama[3]

[1]TU Dortmund University, [2]University of Bonn, [3]University of Tübingen

EARLI

Self-regulated learning at different levels of granularity: From single lessons to entire lectures

Thessaloniki – 24.08.2023

FAIR

Funded by:

Ministry of Culture and Science
of the State of
North Rhine-Westphalia

# Why Self-Testing in Higher Education?

## The problem

- A significant proportion of students are plagued by low performance or high drop out rates (Benden & Lauermann, 2022; Chen & Soldner, 2013; Heublein & Schmelzer, 2018)
- This goes along with significant personal and societal costs (Faas et al., 2018; OECD, 2019)
- Academic performance is one important reason for student drop-out (Yair et al, 2020)
- Study behavior and performance are difficult to improve (Jaggars & Xu, 2016; Oreopoulos & Petronijevic, 2020; Sussan & Son, 2014; Susser & McCabe, 2013)

# Why Self-Testing in Higher Education?

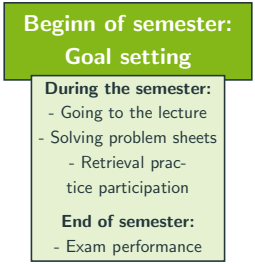
## The problem

- A significant proportion of students are plagued by low performance or high drop out rates (Benden & Lauermann, 2022; Chen & Soldner, 2013; Heublein & Schmelzer, 2018)
- This goes along with significant personal and societal costs (Faas et al., 2018; OECD, 2019)
- Academic performance is one important reason for student drop-out (Yair et al, 2020)
- Study behavior and performance are difficult to improve (Jaggars & Xu, 2016; Oreopoulos & Petronijevic, 2020; Sussan & Son, 2014; Susser & McCabe, 2013)

## Promising learning strategy

- Retrieval practice through self-testing is one of the best-researched learning techniques in the educational sciences (Yang et al., 2021)
- Digital retrieval practice opportunities can help to regulate learning during the semester
- How do initial self-regulated learning (SRL) goals predict SRL-behaviors?
- What is the relative importance of SRL-goals on SRL-behavior?

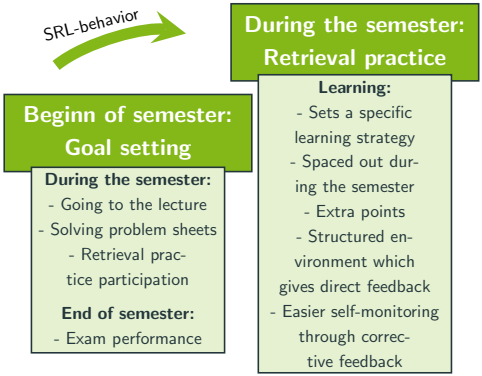
# Do initial SRL-goals predict SRL-behaviors?




Forethought phase for the semester:

- What goals do students set at the beginning of the semester?

# Do initial SRL-goals predict SRL-behaviors?



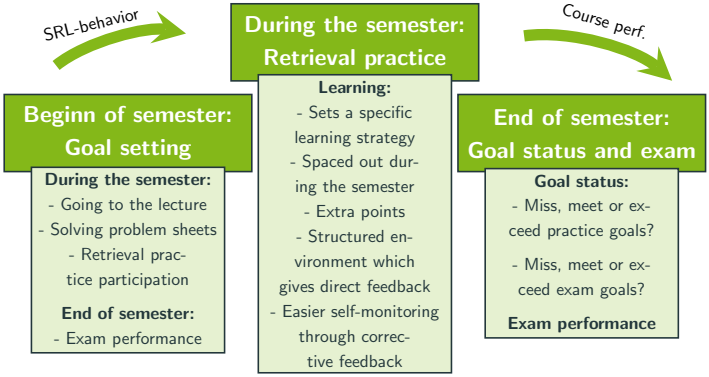

Forethought phase for the semester:

- What goals do students set at the beginning of the semester?

Performance phase during the semester:

- Who is more likely to practice?

# Do initial SRL-goals predict SRL-behaviors?




Forethought phase for the semester:

- What goals do students set at the beginning of the semester?

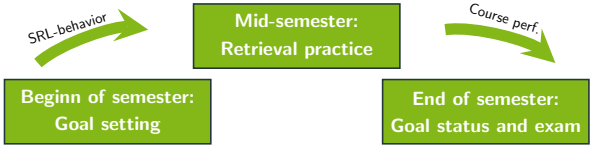
Performance phase during the semester:

- Who is more likely to practice?

Reflection phase at the end of the semester:

- What is students' goal status?

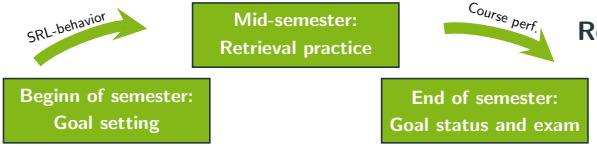
# Do initial SRL-goals predict SRL-behaviors and course performance?




**Retrieval practice on course performance**

- Spaced retrieval practice relates to increased exam performance (Mefferd & Bernacki, 2022; Schwerter & Brahm, 2022)
- Particularly low-achieving students benefit (Förster et al., 2018; Schwerter et al., 2022)
- But: Frequency of retrieval practice is low (Dunlosky & Rawson, 2015, Sussan & Son, 2014; Susser & McCabe, 2013)

# Do initial SRL-goals predict SRL-behaviors and course performance?



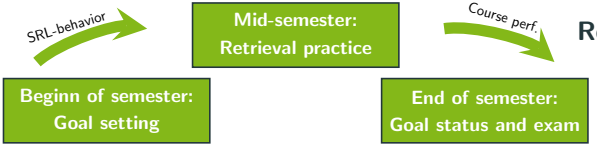

**Goal setting on course performance**

- Final grade goals increase grades (van Lent & Souverijn, 2020)
- Meeting goals can lead to a sense of self-achievement making pursuing goals worthwhile (Gómez-Miñambres, 2012)

**Retrieval practice on course performance**

- Spaced retrieval practice relates to increased exam performance (Mefferd & Bernacki, 2022; Schwerter & Brahm, 2022)
- Particularly low-achieving students benefit (Förster et al., 2018; Schwerter et al., 2022)
- But: Frequency of retrieval practice is low (Dunlosky & Rawson, 2015, Sussan & Son, 2014; Susser & McCabe, 2013)

# Do initial SRL-goals predict SRL-behaviors and course performance?




**Goal setting on course performance**

- Final grade goals increase grades (van Lent & Souverijn, 2020)
- Meeting goals can lead to a sense of self-achievement making pursuing goals worthwhile (Gómez-Miñambres, 2012)

**Retrieval practice on course performance**

- Spaced retrieval practice relates to increased exam performance (Mefferd & Bernacki, 2022; Schwerter & Brahm, 2022)
- Particularly low-achieving students benefit (Förster et al., 2018; Schwerter et al., 2022)
- But: Frequency of retrieval practice is low (Dunlosky & Rawson, 2015, Sussan & Son, 2014; Susser & McCabe, 2013)

## What might moderate the effect of goals on practicing?

- Prior achievement and student background information

# What might moderate the effect of goals on practicing?

- Prior achievement and student background information
- Achievement goals (Elliot & McGregor, 2001)
  - Mastery approach: "My goal is to learn as much as I can in this course."
  - Mastery avoidance: "My goal is to avoid learning less than I could in this course."
  - Performance approach: "I strive to do well compared to other students in this course."
  - Performance avoidance: "I strive to avoid being worse than others in this course."

# What might moderate the effect of goals on practicing?

- Prior achievement and student background information
- Achievement goals (Elliot & McGregor, 2001)
  - Mastery approach: "My goal is to learn as much as I can in this course."
  - Mastery avoidance: "My goal is to avoid learning less than I could in this course."
  - Performance approach: "I strive to do well compared to other students in this course."
  - Performance avoidance: "I strive to avoid being worse than others in this course."
  - Mastery goal orientation is positively correlated with individuals' self-regulatory strategies (Pintrich et al., 1994)
  - Performance goal orientation is usually detrimental to SRL (Zimmerman, 2011)

# What might moderate the effect of goals on practicing?

- Prior achievement and student background information
- Achievement goals (Elliot & McGregor, 2001)
- Expectancy-value beliefs (Eccles et al., 1983)
- Personality traits (Goldberg, 1993)
- Time preference (Frederick & Loewenstein, 2002)

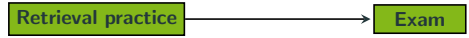
# What might moderate the effect of goals on practicing?

- Prior achievement and student background information
- Achievement goals (Elliot & McGregor, 2001)
- Expectancy-value beliefs (Eccles et al., 1983)
- Personality traits (Goldberg, 1993)
- Time preference (Frederick & Loewenstein, 2002)

→ Integrating different theoretical perspectives (Byrnes & Miller, 2007)

- ML methods: What is the relative importance? (Van Lissa et al., 2023)
- ML methods: Which (non)linear interactions are relevant?

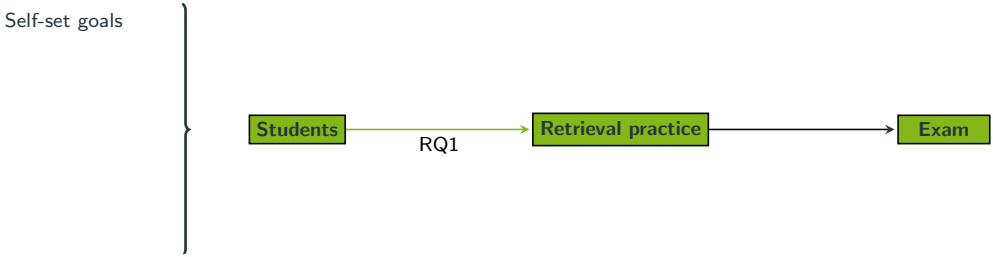
# Present study



# Present study


RQ1: Who takes more practice tests?

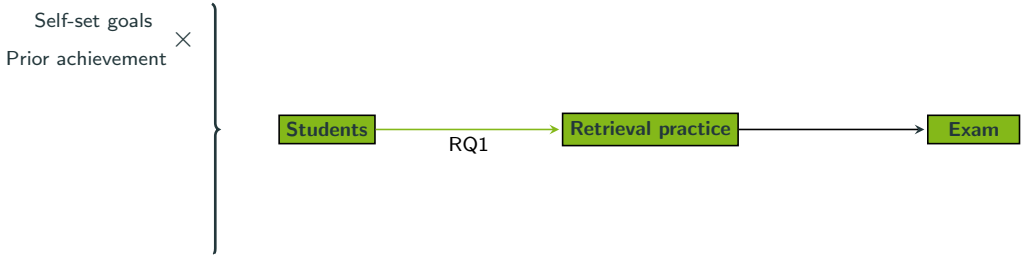
# Present study




RQ1: Who takes more practice tests?

- RQ1a: Do self-set goals predict practice behavior?

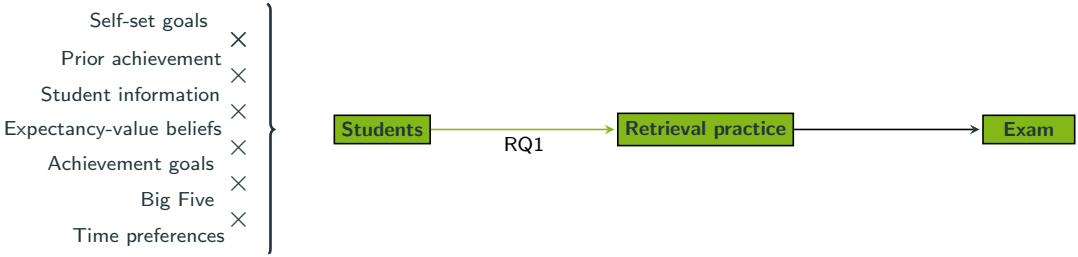
# Present study




RQ1: Who takes more practice tests?

- RQ1a: Do self-set goals predict practice behavior?
- RQ1b: Is the effect moderated by prior-achievement?

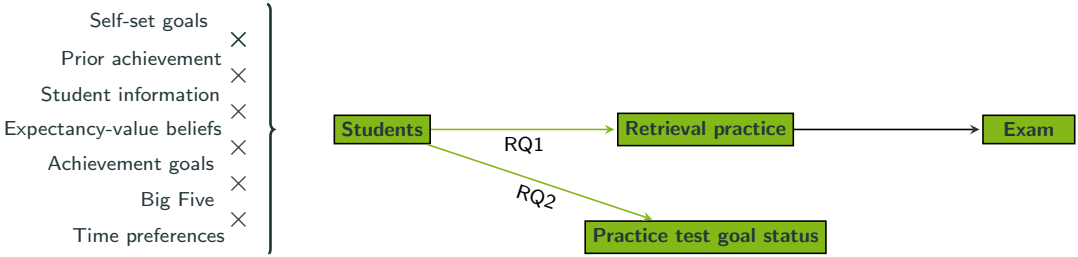
# Present study




RQ1: Who takes more practice tests?

- RQ1a: Do self-set goals predict practice behavior?
- RQ1b: Is the effect moderated by prior-achievement?
- RQ1c: What is the relative importance of goal setting?
  Which other (non-linear) interactions are relevant?

# Present study




RQ1: Who takes more practice tests?

- RQ1a: Do self-set goals predict practice behavior?
- RQ1b: Is the effect moderated by prior-achievement?
- RQ1c: What is the relative importance of goal setting?
  Which other (non-linear) interactions are relevant?

RQ2: What predicts retrieval practice goal status?

## Sample

- Economics & business administration students (N=416, 54% female)
- Winter term 2019/20 at a southwest German university
- $1^{st}$ semester gateway math course (69% first-semester students)
- 3 online practice tests with incentive during the semester & corrective feedback
- Practice tests were available to retake without incentive
- Survey assessing participant information, prior achievement, motivational beliefs, personality traits, time preferences, and self-set course goals

# Measures

- Student information
  - Gender, number of semester, field of study, minor/major, work to finance to study
- Prior achievement
  - HS GPA, advanced math in HS, last math grade in HS
- Achievement goals (Elliot and Murayama, 2008; $\alpha = .64 - .92$)
  - Mastery approach & avoidance; Performance approach & avoidance
- Expectancy value beliefs (Gaspard et al., 2017; $\alpha = .71 - .88$)
  - Self-concept; intrinsic, attainment, & utility value; cost
- Big five (Schupp and Gerlitz, 2014; $\alpha = .65 - .82$)
  - Conscientiousness, Extraversion, Agreeableness, Openness, Neuroticism
- Time preferences (Frederick and Loewenstein, 2002)
  - Risk, Discount factor, Present bias

# Practice, Goals, and Goal status

| Variables | Explanation | N | Mean | SD |
|---|---|---|---|---|
| **Practice test attempts (PTA)** | | | | |
| PTA with incentive | Number of practice tests with incentive | 416 | 2.36 | 0.99 |
| PTA without incentive | Number of practice tests without incentive | 416 | 0.36 | 0.68 |
| Total PTA | Number of practice tests with and without incentive | 416 | 2.72 | 1.26 |
| **Self-set course goals** | | | | |
| PTA goal with incentive | Number of practice test with incentive | 310 | 2.78 | 0.51 |
| PTA goal without incentive | Intention to self-test without incentive (No vs. Yes) | 310 | 0.75 | 0.44 |
| **Goals status / achievement** | | | | |
| PTA goal status with incentive | Miss, meet or exceed practice test attempts goal | 310 | -0.10 | 0.57 |
| PTA goal status w/o incentive | Miss, meet or exceed intention to self-test without incentive | 310 | -0.48 | 0.60 |

Note: The practice variables were observed on the online learning platform of the university. The goals were self-reported in the beginning of the semester. Goal status: miss = -1, meet = 0 & exceed = 1.

# RQ1a: Do self-set goals predict practice behavior?

| | Full sample | | |
|---|---|---|---|
| | PTA in total | PTA with incentive | PTA w/o incentive |
| (Intercept) | 1.438* | 1.177$^{+}$ | 0.261 |
| | (0.682) | (0.604) | (0.285) |
| PTA goal with incentive | 0.502*** | 0.472*** | 0.031 |
| | (0.140) | (0.128) | (0.065) |
| PTA goal w/o incentive | −0.143 | −0.048 | −0.096 |
| | (0.148) | (0.126) | (0.076) |
| Exam performance goal | 0.273** | 0.251*** | 0.022 |
| | (0.083) | (0.066) | (0.042) |
| R2 | 0.135 | 0.168 | 0.013 |
| Adj.R2 | 0.126 | 0.160 | 0.003 |
| N | 416 | 416 | 416 |

PTA = Practice tests attempts; $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

# RQ1a: Do self-set goals predict practice behavior?

| | Full sample | | |
|---|---|---|---|
| | PTA in total | PTA with incentive | PTA w/o incentive |
| (Intercept) | 1.438* | 1.177$^{+}$ | 0.261 |
| | (0.682) | (0.604) | (0.285) |
| PTA goal with incentive | 0.502*** | 0.472*** | 0.031 |
| | (0.140) | (0.128) | (0.065) |
| PTA goal w/o incentive | −0.143 | −0.048 | −0.096 |
| | (0.148) | (0.126) | (0.076) |
| Exam performance goal | 0.273** | 0.251*** | 0.022 |
| | (0.083) | (0.066) | (0.042) |
| R2 | 0.135 | 0.168 | 0.013 |
| Adj.R2 | 0.126 | 0.160 | 0.003 |
| N | 416 | 416 | 416 |

PTA = Practice tests attempts; $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

$\Rightarrow$ Setting higher goals is associated with having more practice test attempts (with incentive)

# RQ1a: Do self-set goals predict practice behavior?

| | Full sample | | |
|---|---|---|---|
| | PTA in total | PTA with incentive | PTA w/o incentive |
| (Intercept) | 1.438* | 1.177$^{+}$ | 0.261 |
| | (0.682) | (0.604) | (0.285) |
| PTA goal with incentive | 0.502*** | 0.472*** | 0.031 |
| | (0.140) | (0.128) | (0.065) |
| PTA goal w/o incentive | −0.143 | −0.048 | −0.096 |
| | (0.148) | (0.126) | (0.076) |
| Exam performance goal | 0.273** | 0.251*** | 0.022 |
| | (0.083) | (0.066) | (0.042) |
| R2 | 0.135 | 0.168 | 0.013 |
| Adj.R2 | 0.126 | 0.160 | 0.003 |
| N | 416 | 416 | 416 |

PTA = Practice tests attempts; $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

$\Rightarrow$ Setting higher goals is associated with having more practice test attempts (with incentive)

- Is this effect driven by high-achieving students? (RQ1b)

# RQ1b: Interacting goals with prior achievement

| | Practice tests attempts | | |
|---|---|---|---|
| | Total | Total | Total |
| PTA goal with incentive | 0.425** | 0.503* | 0.421** |
| | (0.133) | (0.199) | (0.132) |
| PTA goal w/o incentive | −0.177 | −0.144 | −0.204 |
| | (0.136) | (0.148) | (0.142) |
| Exam performance goal | 0.141+ | 0.272** | 0.140 |
| | (0.084) | (0.082) | (0.091) |
| HS GPA | 0.534+ | | |
| | (0.308) | | |
| PTA goal w.i. × HS GPA | −0.076 | | |
| | (0.112) | | |
| Advanced math in HS | | 0.029 | |
| | | (0.728) | |
| PTA goal w.i. × Advanced math in HS | | −0.008 | |
| | | (0.258) | |
| Last math grade in HS | | | 0.484 |
| | | | (0.328) |
| PTA goal w.i. × Last math grade in HS | | | −0.089 |
| | | | (0.117) |
| Adj.R2 | 0.184 | 0.124 | 0.153 |
| N | 416 | 416 | 416 |

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

# RQ1b: Interacting goals with prior achievement

| | Practice tests attempts | | |
|---|---|---|---|
| | Total | Total | Total |
| PTA goal with incentive | 0.425** | 0.503* | 0.421** |
| | (0.133) | (0.199) | (0.132) |
| PTA goal w/o incentive | −0.177 | −0.144 | −0.204 |
| | (0.136) | (0.148) | (0.142) |
| Exam performance goal | 0.141+ | 0.272** | 0.140 |
| | (0.084) | (0.082) | (0.091) |
| HS GPA | 0.534+ | | |
| | (0.308) | | |
| PTA goal w.i. × HS GPA | −0.076 | | |
| | (0.112) | | |
| Advanced math in HS | | 0.029 | |
| | | (0.728) | |
| PTA goal w.i. × Advanced math in HS | | −0.008 | |
| | | (0.258) | |
| Last math grade in HS | | | 0.484 |
| | | | (0.328) |
| PTA goal w.i. × Last math grade in HS | | | −0.089 |
| | | | (0.117) |
| Adj.R2 | 0.184 | 0.124 | 0.153 |
| N | 416 | 416 | 416 |

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

⇒ The effect of goal setting is not driven by high-achieving students

# RQ1b: Interacting goals with prior achievement

| | Practice tests attempts | | |
|---|---|---|---|
| | Total | Total | Total |
| PTA goal with incentive | 0.425** | 0.503* | 0.421** |
| | (0.133) | (0.199) | (0.132) |
| PTA goal w/o incentive | −0.177 | −0.144 | −0.204 |
| | (0.136) | (0.148) | (0.142) |
| Exam performance goal | 0.141+ | 0.272** | 0.140 |
| | (0.084) | (0.082) | (0.091) |
| HS GPA | 0.534+ | | |
| | (0.308) | | |
| PTA goal w.i. × HS GPA | −0.076 | | |
| | (0.112) | | |
| Advanced math in HS | | 0.029 | |
| | | (0.728) | |
| PTA goal w.i. × Advanced math in HS | | −0.008 | |
| | | (0.258) | |
| Last math grade in HS | | | 0.484 |
| | | | (0.328) |
| PTA goal w.i. × Last math grade in HS | | | −0.089 |
| | | | (0.117) |
| Adj.R2 | 0.184 | 0.124 | 0.153 |
| N | 416 | 416 | 416 |

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

⇒ The effect of goal setting is not driven by high-achieving students

- What predicts practice test attempts? (RQ1c)

## RQ1c – What predicts total practice test attempts?

- Prediction Rule Ensembles (Fokkema & Strobl, 2022)
- Tree-based method combined with LASSO to find the most important (non-linear) interactions of variables

# RQ1c – What predicts total practice test attempts?

- Prediction Rule Ensembles (Fokkema & Strobl, 2022)
- Tree-based method combined with LASSO to find the most important (non-linear) interactions of variables

**Variable importance**

| | Variables | Import. |
|---|---|---|
| 1 | HS GPA | 0.06 |
| 2 | Number of semester | 0.04 |
| 3 | Mast. approach | 0.03 |
| 4 | Conscientiousness | 0.03 |
| 5 | PTA goal with incen. | 0.02 |
| 6 | Perf. approach | 0.01 |
| 7 | Work to study | 0.01 |
| 8 | Non-major | 0.01 |
| 9 | Perf. avoidance | 0.01 |

Import. = Importance

**Prediction rules**

| Rules | Coef. |
|---|---|
| Mast. approach $\leq 7$ & HS GPA $\leq 3.5$ | -0.17 |
| Conscientiousness $\leq 8$ & HS GPA $\leq 3.7$ | -0.10 |
| PTA goal with incentive $= 3$ & Num. semester $= 1$ | 0.09 |
| Perf. approach $> 4$ & Num. semester $\leq 3$ | 0.05 |
| Num. semester $= 1$ & Perf. avoidance $> 2$ | 0.04 |
| Non-major $= 0$ & Work to study $= 0$ | 0.05 |

PTA = Practice test attempts

# RQ1c – What predicts total practice test attempts?

- Prediction Rule Ensembles (Fokkema & Strobl, 2022)
- Tree-based method combined with LASSO to find the most important (non-linear) interactions of variables

**Variable importance**

| | Variables | Import. |
|---|---|---|
| 1 | HS GPA | 0.06 |
| 2 | Number of semester | 0.04 |
| 3 | Mast. approach | 0.03 |
| 4 | Conscientiousness | 0.03 |
| 5 | PTA goal with incen. | 0.02 |
| 6 | Perf. approach | 0.01 |
| 7 | Work to study | 0.01 |
| 8 | Non-major | 0.01 |
| 9 | Perf. avoidance | 0.01 |

Import. = Importance

**Prediction rules**

| Rules | Coef. |
|---|---|
| Mast. approach $\leq 7$ & HS GPA $\leq 3.5$ | -0.17 |
| Conscientiousness $\leq 8$ & HS GPA $\leq 3.7$ | -0.10 |
| PTA goal with incentive $= 3$ & Num. semester $= 1$ | 0.09 |
| Perf. approach $> 4$ & Num. semester $\leq 3$ | 0.05 |
| Num. semester $= 1$ & Perf. avoidance $> 2$ | 0.04 |
| Non-major $= 0$ & Work to study $= 0$ | 0.05 |

PTA = Practice test attempts

$\Rightarrow$ Especially first semester students with a high goal had more attempts

# RQ1c – What predicts total practice test attempts?

- Prediction Rule Ensembles (Fokkema & Strobl, 2022)
- Tree-based method combined with LASSO to find the most important (non-linear) interactions of variables

**Variable importance**

| | Variables | Import. |
|---|---|---|
| 1 | HS GPA | 0.06 |
| 2 | Number of semester | 0.04 |
| 3 | Mast. approach | 0.03 |
| 4 | Conscientiousness | 0.03 |
| 5 | PTA goal with incen. | 0.02 |
| 6 | Perf. approach | 0.01 |
| 7 | Work to study | 0.01 |
| 8 | Non-major | 0.01 |
| 9 | Perf. avoidance | 0.01 |

Import. = Importance

**Prediction rules**

| Rules | Coef. |
|---|---|
| Mast. approach $\leq 7$ & HS GPA $\leq 3.5$ | -0.17 |
| Conscientiousness $\leq 8$ & HS GPA $\leq 3.7$ | -0.10 |
| PTA goal with incentive $= 3$ & Num. semester $= 1$ | 0.09 |
| Perf. approach $> 4$ & Num. semester $\leq 3$ | 0.05 |
| Num. semester $= 1$ & Perf. avoidance $> 2$ | 0.04 |
| Non-major $= 0$ & Work to study $= 0$ | 0.05 |

PTA = Practice test attempts

$\Rightarrow$ Especially first semester students with a high goal had more attempts

- Interactions of HS GPA with personality and achievement goals

# RQ1c – What predicts practice test attempts?

**Figure 1:** Interaction between mastery approach & HS GPA

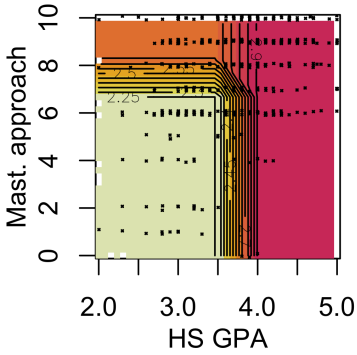
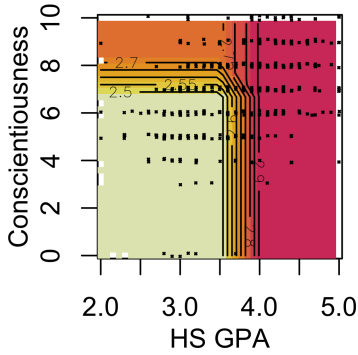



**Figure 2:** Interaction between conscientiousness & HS GPA



# RQ1c – What predicts practice test attempts?

**Figure 1:** Interaction between mastery approach & HS GPA

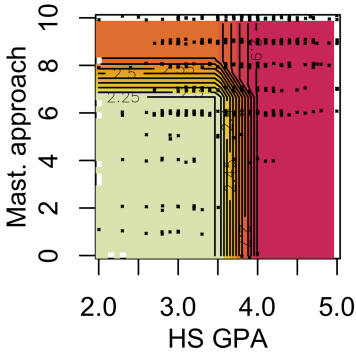
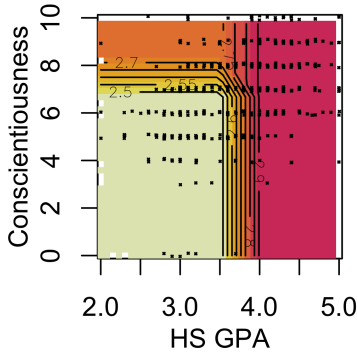



**Figure 2:** Interaction between conscientiousness & HS GPA




⇒ High-performing students "have figured it out"

# RQ1c – What predicts practice test attempts?

**Figure 1:** Interaction between mastery approach & HS GPA

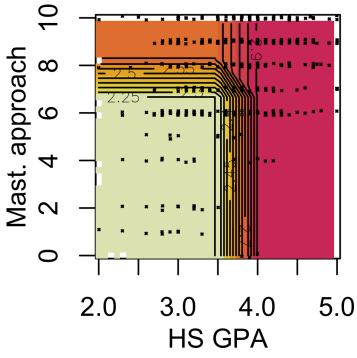
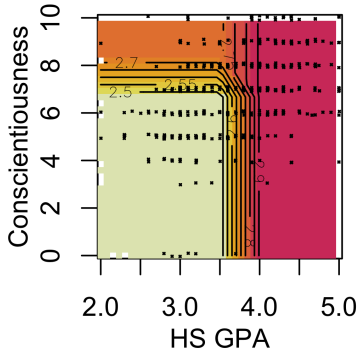



**Figure 2:** Interaction between conscientiousness & HS GPA




⇒ High-performing students "have figured it out"

- High mastery approach and conscientiousness help compensate for lower HS GPA

# RQ2 – What predicts the practice test attempts goal status?

**Variable importance**

| | Variables | Import. |
|---|---|---|
| 1 | HS GPA | 0.21 |
| 2 | PTA goal with incen. | 0.12 |
| 3 | Perf. approach | 0.12 |
| 4 | Conscientiousness | 0.08 |
| 5 | Mast. avoidance | 0.05 |
| 6 | Number of semester | 0.04 |
| 7 | Mast. approach | 0.02 |
| 8 | Non-major | 0.02 |
| 9 | Extraversion | 0.01 |

Import. = Importance

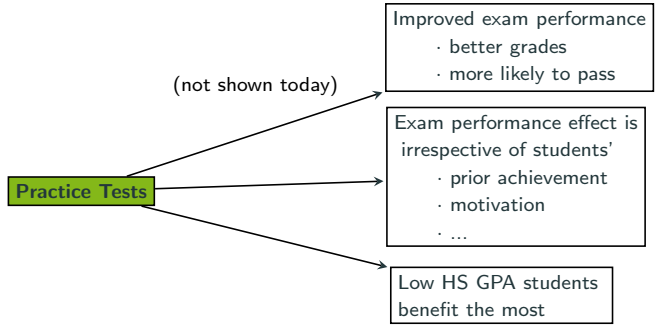
**Prediction rules**

| Rules | Import. | Coef. |
|---|---|---|
| Perf. approach $>$ 3 & PTA goal with incen. $\leq$ 2 | 0.24 | 0.69 |
| Conscientiousness $\leq$ 7 & HS GPA $\leq$ 3.9 | 0.12 | -0.23 |
| HS GPA $\leq$ 3.5 & Mast. avoidance $\leq$ 6 | 0.10 | -0.24 |
| HS GPA $>$ 3.2 & Number of semester $\leq$ 1 | 0.08 | 0.18 |
| HS GPA $\leq$ 3.7 & Mast. approach $\leq$ 9 | 0.05 | -0.09 |
| HS GPA $>$ 3.5 & Non-major $\leq$ 0 | 0.04 | 0.08 |
| Conscientiousness $\leq$ 7 & HS GPA $\leq$ 3.5 | 0.03 | 0.06 |
| HS GPA $>$ 3.5 & Extraversion $>$ 4 | 0.01 | 0.02 |
| HS GPA $\leq$ 3.7 & Conscientiousness $\leq$ 7 | 0.01 | 0.01 |

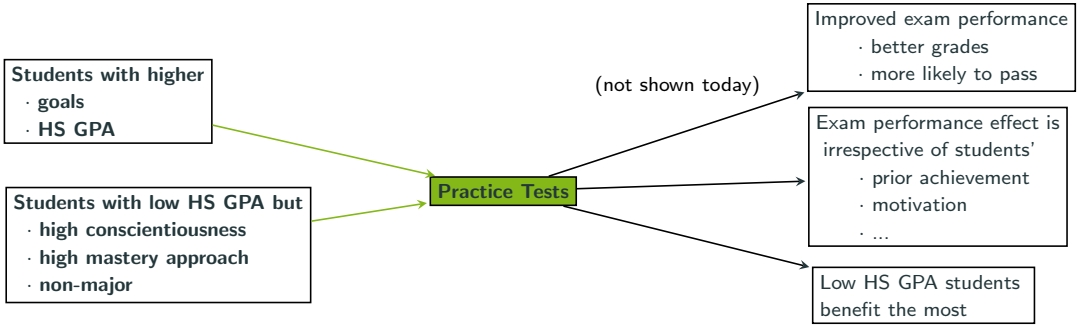
PTA w.i. = Practice test attempts With incentive

- Lower goals but high performance approach increase goal status
- For students with a high HS GPA it does not matter
- Students with low HS GPA “need” high cons., mast. approach., and mast. avoidance

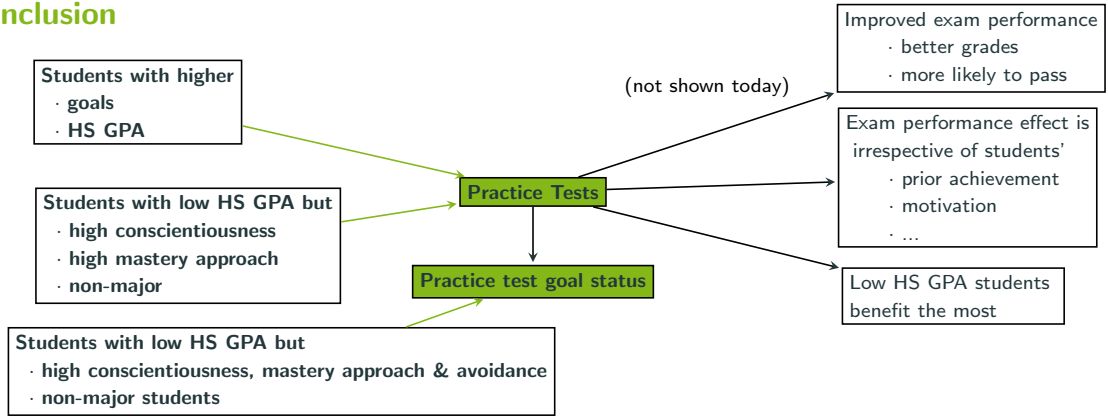
# Conclusion


Improved exam performance
· better grades
· more likely to pass
(not shown today)
Exam performance effect is irrespective of students'
· prior achievement
· motivation
· ...
Practice Tests
Low HS GPA students benefit the most

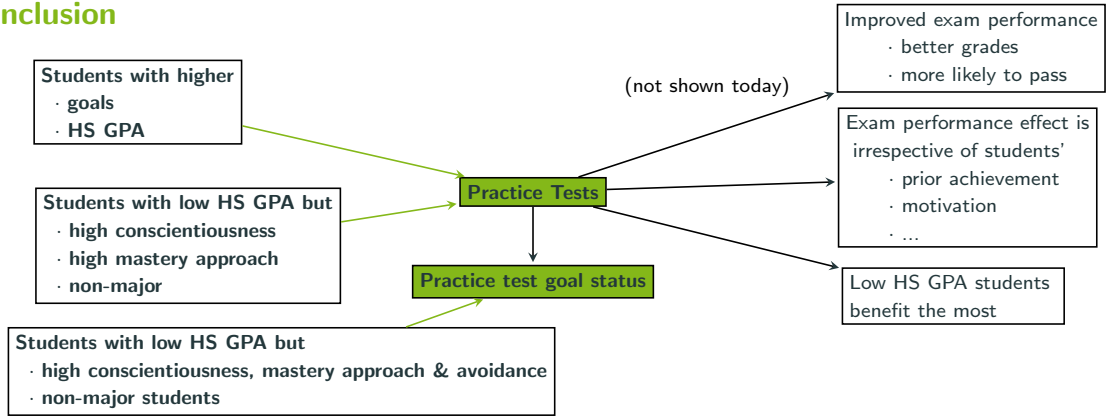
# Conclusion


Students with higher
· goals
· HS GPA
Students with low HS GPA but
· high conscientiousness
· high mastery approach
· non-major
Practice Tests
(not shown today)
Improved exam performance
· better grades
· more likely to pass
Exam performance effect is irrespective of students'
· prior achievement
· motivation
· ...
Low HS GPA students benefit the most

# Conclusion


Students with higher
· goals
· HS GPA
Students with low HS GPA but
· high conscientiousness
· high mastery approach
· non-major
Students with low HS GPA but
· high conscientiousness, mastery approach & avoidance
· non-major students
Practice Tests
Practice test goal status
(not shown today)
Improved exam performance
· better grades
· more likely to pass
Exam performance effect is irrespective of students'
· prior achievement
· motivation
· ...
Low HS GPA students benefit the most

# Conclusion

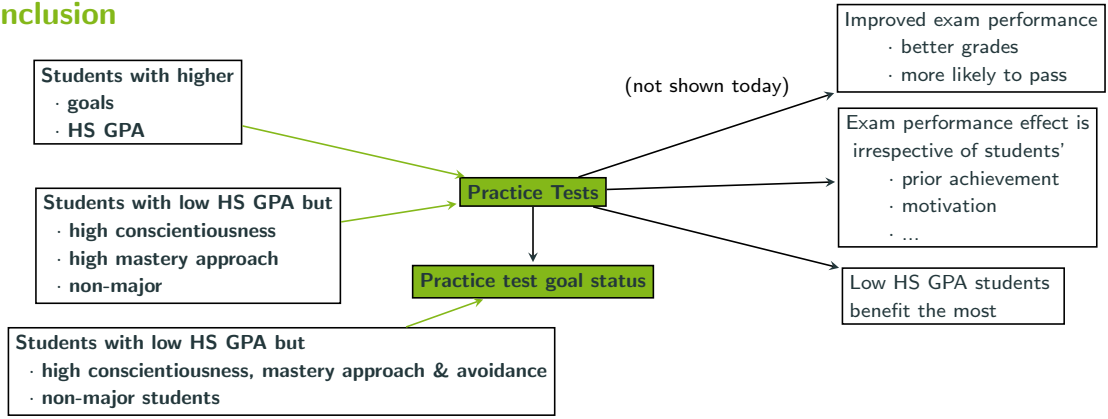



⇒ **Higher goals are related to more practice test attempts (with incentive)**

⇒ **Students with high prior achievement "have figured it out"**

⇒ **Conscientiousness & mastery approach can help compensate for low prior achievement**

# Conclusion




⇒ **Higher goals are related to more practice test attempts (with incentive)**

⇒ **Students with high prior achievement "have figured it out"**

⇒ **Conscientiousness & mastery approach can help compensate for low prior achievement**

From prediction to intervention:

- Intervention on student goal setting and achievement goals

⇒ Increase practice test attempts and thus exam performance?

# Thank you!

**Dr. Jakob Schwerter**
jakob.schwerter@tu-dortmund.de
https://www.researchgate.net/profile/Jakob-Schwerter

- Schwerter, J., Dimpfl, T., Bleher, J., & Murayama, K. (2022). Benefits of additional online practice opportunities in higher education. *Internet and Higher Education, 53*, 100834. https://doi.org/https://doi.org/10.1016/j.iheduc.2021.100834
- Schwerter, J., & Brahm, T. (2022). Voluntary e-learning exercises support students in mastering statistics. *Preprint at SocArXiv*. https://doi.org/10.31235/osf.io/8p2xh
- Schwerter, J., Wortha, F., & Gerjets, P. (2022). E-Learning with multiple-try-feedback: Can hints foster students' achievement during the semester? *Educational Technology Research and Development, 70*, 713-736. https://doi.org/10.1007/s11423-022-10105-z

Funded by:
Ministry of Culture and Science of the State of North Rhine-Westphalia

# Scales

| Variables | Explanation | N | Mean | SD | $\alpha$ |
|---|---|---|---|---|---|
| **Expectancy value beliefs** | | | | | |
| Self-concept | Math is easy for me | 318 | 2.65 | 0.64 | 0.86 |
| Intrinsic value/Dispositional Interest | Math is fun for me | 316 | 2.69 | 0.61 | 0.87 |
| Attainment value | It is personally important to me to master mathematical content | 315 | 2.95 | 0.56 | 0.71 |
| Utility value | For my professional future it will pay off to be good at math | 314 | 3.47 | 0.55 | 0.88 |
| Cost | Dealing with math costs me a lot of energy | 315 | 2.39 | 0.55 | 0.75 |
| **Achievement goals** | | | | | |
| Mastery approach | My goal is to learn as much as I can in this course | 308 | 6.08 | 0.79 | 0.64 |
| Mastery avoidance | My goal is to avoid learning less than I could in this course | 308 | 5.60 | 1.01 | 0.71 |
| Performance approach | I strive to do well compared to other students in this course | 303 | 4.81 | 1.56 | 0.87 |
| Performance avoidance | I strive to avoid being worse than others in this course | 210 | 5.01 | 1.59 | 0.92 |
| **Big five** | | | | | |
| Conscientiousness | ... works thoroughly | 313 | 4.90 | 1.08 | 0.65 |
| Extraversion | ... is communicative, talkative | 313 | 4.93 | 1.30 | 0.82 |
| Agreeableness | ... can forgive | 313 | 5.51 | 1.05 | 0.62 |
| Openness | ... is original, introduces new ideas | 311 | 4.91 | 1.17 | 0.65 |
| Neuroticism | ... is slightly nervous | 313 | 4.38 | 1.26 | 0.68 |

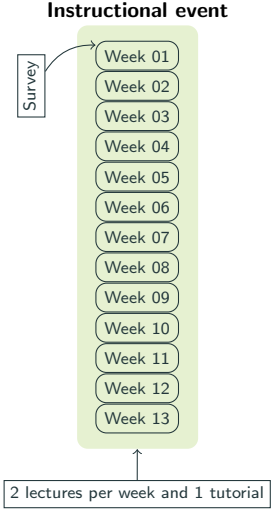
# Math gap

## Why is mathematics a special case?

- Self-testing in mathematics is conceptually different from fact recall:
- Solution steps $\neq$ solutions $\Rightarrow$ "Derivative of the function $f(x) = 2 \cdot e^{4 \cdot x} + 4x + 2$"
- Little evidence for higher education math courses
  (Carvalho et al., 2022; Förster et al., 2018; Wong et al., 2017; Yeo & Fazio, 2019)
  - Laboratory studies with exercises involving the Poisson distribution (Yeo & Fazio, 2019)
  - Online quizzes in a statistics course (Förster et al., 2018)

? Unclear how, when, and for whom practice testing in mathematics is required

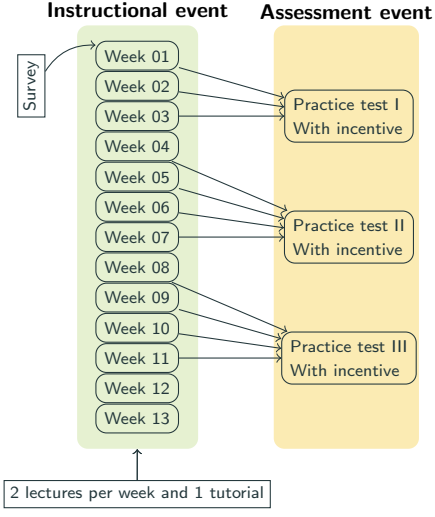
# Research Design & Sample




## Sample and Context

- Economics/business administration students (N=188, 59% female)
- 1st semester gateway math course

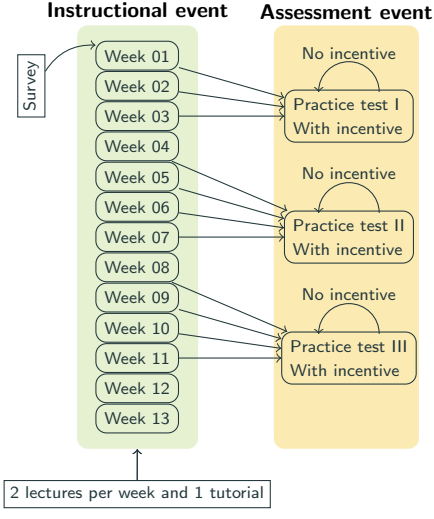
# Research Design & Sample




## Sample and Context

- Economics/business administration students (N=188, 59% female)
- 1st semester gateway math course

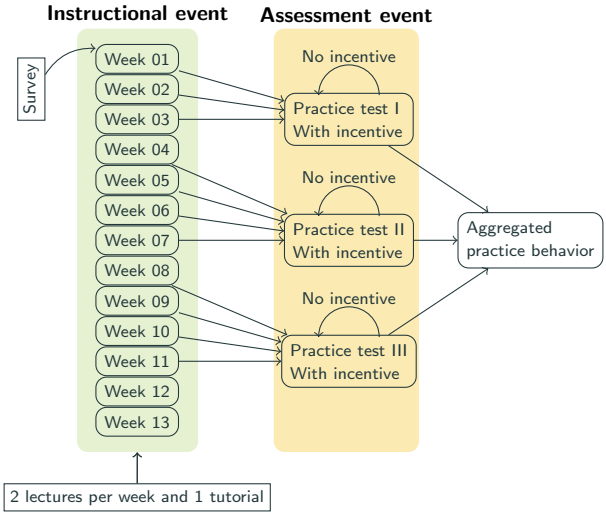
# Research Design & Sample




## Sample and Context

- Economics/business administration students (N=188, 59% female)
- 1st semester gateway math course
- Practice tests: Content from the previous 3-to-4 weeks
- Immediate corrective feedback

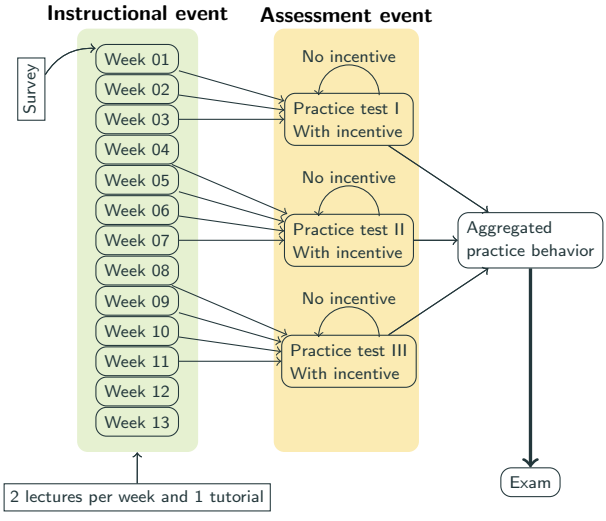
# Research Design & Sample




## Sample and Context

- Economics/business administration students (N=188, 59% female)
- 1st semester gateway math course
- Practice tests: Content from the previous 3-to-4 weeks
- Immediate corrective feedback

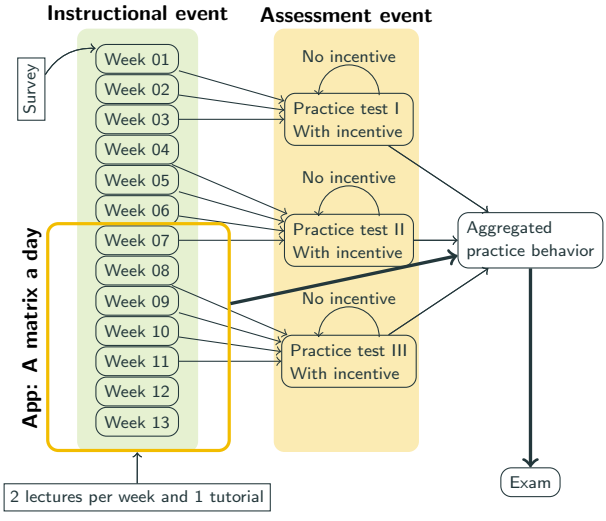
# Research Design & Sample




## Sample and Context

- Economics/business administration students (N=188, 59% female)
- 1st semester gateway math course
- Practice tests: Content from the previous 3-to-4 weeks
- Immediate corrective feedback

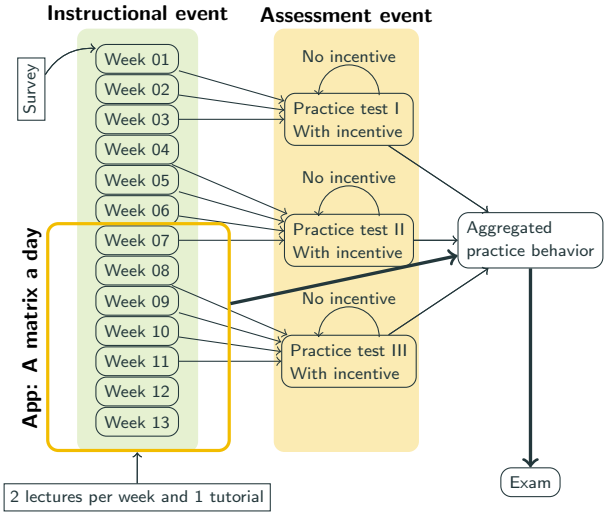
# Research Design & Sample




## Sample and Context

- Economics/business administration students (N=188, 59% female)
- 1st semester gateway math course
- Practice tests: Content from the previous 3-to-4 weeks
- Immediate corrective feedback

# Research Design & Sample




## Sample and Context

- Economics/business administration students (N=188, 59% female)
- 1st semester gateway math course
- Practice tests: Content from the previous 3-to-4 weeks
- Immediate corrective feedback
- Explanatory variables:
  - Practice test attempts
  - Practice test performance
  - App submissions
  - App performance
- Outcome: Points in final exam

# Follow-up study

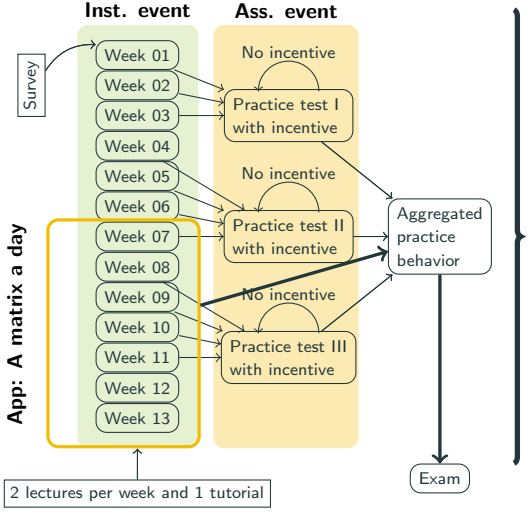
**Practice makes perfect: A deeper dive into the motivational and cognitive underpinnings of retrieval practice**

- Does retrieval practice enable knowledge transfer (in mathematics)?

# Follow-up study

**Practice makes perfect: A deeper dive into the motivational and cognitive underpinnings of the self-testing effect**

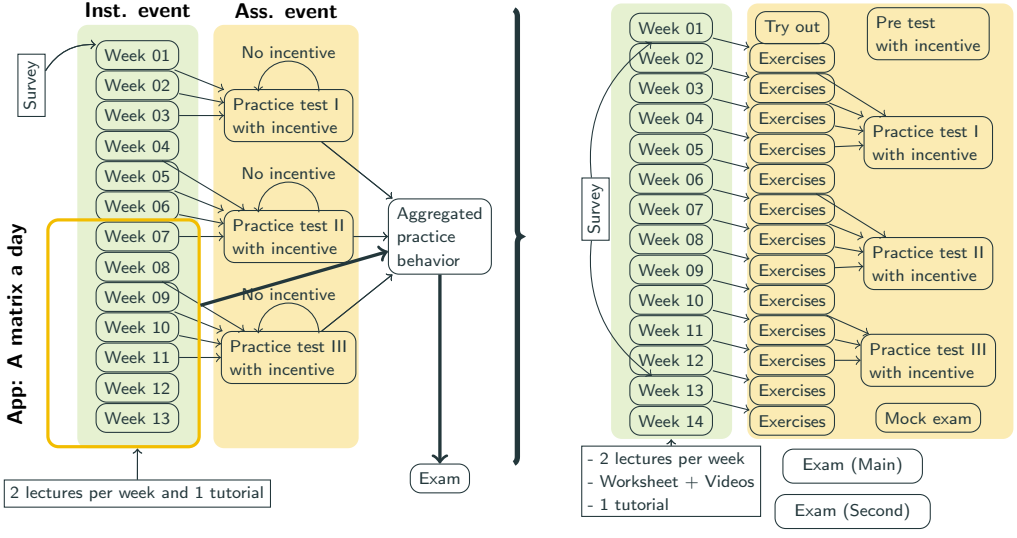
- Does retrieval practice enable knowledge transfer (in mathematics)?



# Follow-up study

**Practice makes perfect: A deeper dive into the motivational and cognitive underpinnings of the self-testing effect**

- Does retrieval practice enable knowledge transfer (in mathematics)?



## **Table 1:** Exam performance

| | Full sample | | | Took exam | | | Schwerter et al. | | |
|---|---|---|---|---|---|---|---|---|---|
| | exam | grade | passed | exam | grade | passed | exam | grade | passed |
| (Intercept) | $-1.881^{***}$ | $0.508^{***}$ | $-0.119^{*}$ | $-1.714^{***}$ | 0.250 | $-0.082$ | $-2.402^{***}$ | $-0.557$ | $-0.304^{+}$ |
| | (0.142) | (0.105) | (0.051) | (0.239) | (0.266) | (0.115) | (0.324) | (0.408) | (0.168) |
| Practice test attempts | $0.244^{***}$ | $0.214^{***}$ | $0.127^{***}$ | $0.147^{**}$ | $0.172^{*}$ | $0.093^{***}$ | $0.226^{**}$ | $0.284^{**}$ | $0.110^{***}$ |
| | (0.045) | (0.051) | (0.022) | (0.056) | (0.072) | (0.025) | (0.073) | (0.097) | (0.030) |
| Practice test performance | $0.013^{***}$ | $0.014^{***}$ | $0.005^{***}$ | $0.017^{***}$ | $0.022^{***}$ | $0.006^{***}$ | $0.022^{***}$ | $0.029^{***}$ | $0.009^{***}$ |
| | (0.002) | (0.003) | (0.001) | (0.003) | (0.004) | (0.002) | (0.004) | (0.005) | (0.002) |
| submissions.ama | $-0.000$ | 0.001 | $-0.002$ | 0.004 | 0.004 | $-0.001$ | 0.007 | 0.007 | 0.001 |
| | (0.008) | (0.010) | (0.003) | (0.009) | (0.011) | (0.003) | (0.011) | (0.014) | (0.002) |
| percentages.ama | $0.006^{***}$ | $0.008^{***}$ | $0.002^{**}$ | $0.005^{**}$ | $0.005^{*}$ | $0.002^{*}$ | $0.004^{*}$ | $0.005^{+}$ | 0.001 |
| | (0.002) | (0.002) | (0.001) | (0.002) | (0.002) | (0.001) | (0.002) | (0.003) | (0.001) |
| R2 | 0.383 | 0.313 | 0.307 | 0.247 | 0.237 | 0.183 | 0.230 | 0.218 | 0.173 |
| Adj.R2 | 0.377 | 0.306 | 0.300 | 0.236 | 0.226 | 0.171 | 0.213 | 0.201 | 0.155 |
| BIC (null) | $-176$ | $-132$ | $-128$ | $-57$ | $-53$ | $-34$ | $-28$ | $-25$ | $-15$ |
| AIC (null) | 216 | 369 | $-378$ | 163 | 295 | $-235$ | 115 | 222 | $-168$ |
| N | 416 | 416 | 416 | 281 | 281 | 281 | 188 | 188 | 188 |
| Number of parameters | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

Note: Full sample: Multiple imputation for all students. Took exam: Multiple imputation only for students who wrote the exam. Schwerter et al.: Complete cases sample from the Schwerter et al. (2022).

# Regression result summary

## How effective are the additional self-testing opportunities?

- Self-testing with the practice tests improved students' exam scores by about 5 points (of 90)
- Self-testing effect remained significant but decreased to 2.5 points
- Use of the daily self-testing app only significant in simple OLS regression
- Risk-averse students, those who planned repeated practice, and students with a higher math self-concept were most likely to benefit from self-testing
- Students' gender, achievement goals, and personality traits did not contribute to differential practice effects
- Only one significant interaction emerged for self-testing via the daily app
- Higher open-mindedness corresponded to greater benefits from self-testing

# Scales

| Variables | Explanation | N | Mean | SD |
|---|---|---|---|---|
| **Individual characteristics** | | | | |
| Female | Female vs Male | 389 | 0.54 | 0.50 |
| High school GPA | Numeric value from 1 to 5 | 315 | 3.81 | 0.62 |
| Advanced math in HS | Indicator if yes or no | 304 | 0.81 | 0.40 |
| Last math grade in HS | Numeric value from 1 to 5 | 315 | 3.24 | 1.16 |
| International studies | Indicator if yes or no | 416 | 0.39 | 0.49 |
| Sports degree | Indicator if yes or no | 416 | 0.06 | 0.24 |
| Minor | Indicator if yes or no | 416 | 0.25 | 0.43 |
| Work to study | Indicator if yes or no | 295 | 0.25 | 0.43 |
| Semester | Number of semester | 313 | 1.33 | 1.31 |
| **Present bias preferences** | | | | |
| Risk | Value bigger .5 implies risk aversion | 307 | 0.68 | 0.20 |
| Discount factor | Value below 1 indicates to be impatient | 301 | 1.00 | 0.67 |
| Present bias | Value below 1 makes individuals more impatient when the present is involved | 300 | 1.06 | 0.31 |

# Who benefits from self-testing opportunities?

## Heterogeneity

- Most of the self-testing results are general treatment effects without much heterogeneity analysis
- Notable exceptions:
  - Students with lower prior knowledge benefit most from self-testing
  - Students with higher prior knowledge become less overconfident with self-testing (Cogliano et al., 2019)
  - Missing: Interactions with motivation or personality variables

## OLS and Post-double selection results

| | Dependent variable: Standardized points on final exam | | | |
|---|---|---|---|---|
| | Practice variables only | PDSR LASSO | PDSR Rand. For. | PDSR XGBoost |
| Constant | -2.401*** | -0.824 | -0.078 | -0.195 |
| | (0.328) | (0.847) | (1.256) | (0.753) |
| Practice test attempts | 0.226*** | 0.215*** | 0.203*** | 0.205*** |
| | (0.074) | (0.068) | (0.065) | (0.064) |
| Practice test performance | 0.022*** | 0.010** | 0.010** | 0.010** |
| | (0.004) | (0.004) | (0.004) | (0.004) |
| App submissions | 0.007 | 0.006 | 0.001 | 0.004 |
| | (0.012) | (0.008) | (0.009) | (0.009) |
| App performance | 0.004** | 0.004** | 0.005*** | 0.004** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Additional controls | No | Yes | Yes | Yes |
| Observations | 188 | 188 | 188 | 188 |
| Adjusted $R^2$ | 0.213 | 0.446 | 0.464 | 0.410 |

PDSR: Post-double selection regression $\Rightarrow$ Selecting important variables for the dependent variable and the four practice variables using LASSO, RF, and XGBoost.

# Heterogeneity explaining exam performance (I)

**Table 2:** Variable importance of selected variables

| | Variables | Importance |
|---|---|---|
| 1 | HS GPA | 0.42 |
| 2 | Practice test performance | 0.22 |
| 3 | Math self-concept | 0.18 |
| 4 | HS math grade | 0.16 |
| 5 | Practice tests attempts | 0.11 |
| 6 | App submissions | 0.09 |
| 7 | App performance | 0.08 |

**Table 3:** Prediction rules

| Rules | Coefficient |
|---|---|
| $SC_{Math} \leq 3.25$ & App submissions $\leq 3$ | -0.37 |
| HS GPA $\leq 3.7$ & App performance $\leq 45.71$ | -0.35 |
| PT attempts $\leq 3$ & HS GPA $\leq 3.7$ | -0.34 |
| HS GPA $> 3.3$ & HS math grade $> 2.6$ | 0.30 |
| HS GPA $> 4.1$ & PT performance $> 65.33$ | 0.28 |
| HS GPA $> 3.6$ & HS math grade $> 2.6$ | -0.20 |
| HS GPA $> 3.7$ & PT performance $> 78.03$ | 0.20 |
| PT performance $> 56.67$ & $SC_{Math} > 2.67$ | 0.17 |
| PT performance $> 56.67$ & HS math grade $> 3$ | 0.16 |
| HS GPA $> 3.3$ & $SC_{Math} > 2.25$ | 0.16 |
| PT attempts $\leq 3$ & PT performance $\leq 72.41$ | -0.14 |

PT = Practice test

*Fokkema, M., & Strobl, C. (2020). Fitting Prediction Rule Ensembles to Psychological Research Data: An Introduction and Tutorial. *Psychological Methods*, 25(5), 636–652. https://doi.org/10.1037/met0000256

# Heterogeneity explaining exam performance (II)

**Figure 3:** Interaction between self-testing attempts & HS GPA

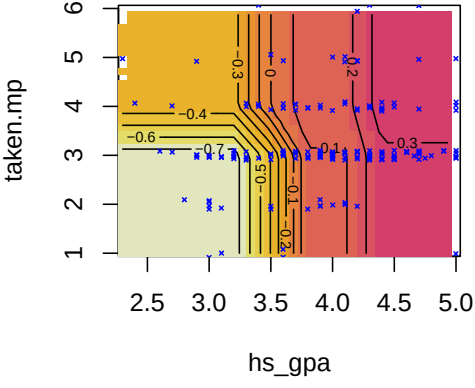
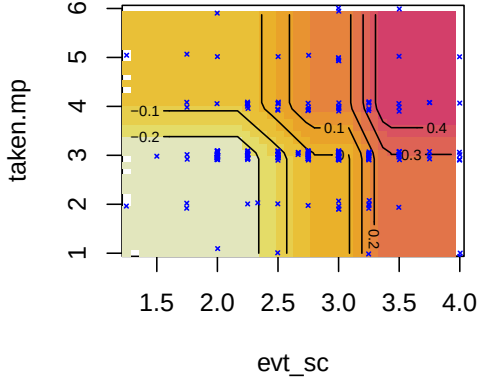
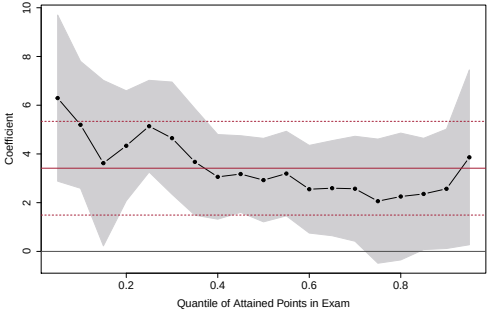



**Figure 4:** Interaction between self-testing attempts & Math self-concept



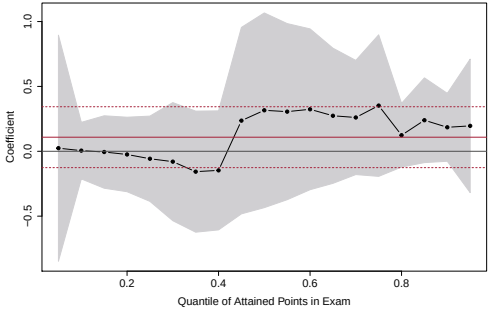
# Quantile regressions

**Figure 5:** Practice test attempts




- Students with fewer exam points would have benefitted more than with more points

**Figure 6:** APP submissions




- Somewhat constant
- Good students would have benefitted from more practice in matrix algebra

## Summary

### Effectiveness of additional online self-testing

Practice tests:

- Effectiveness of self-testing with corrective feedback in a gateway higher education math course
  - Authentic learning environment
  - Higher order learning
  - Parameterized exercises $\rightarrow$ Near content transfer

$\Rightarrow$ Online practice tests are a promising intervention to support students

- Intervention with support for low-performing students

App 'Matrix a Day'

- Course-embedded practice tests were more effective than the daily self-testing app

$\Rightarrow$ The modality of implementation warrants further consideration

# RQ2 – What predicts the practice test attempts goal status?

## Variable importance

| | Variables | Import. |
|---|---|---|
| 1 | HS GPA | 0.16 |
| 2 | PTAwr goal | 0.12 |
| 3 | Perf. approach | 0.12 |
| 4 | Non-major | 0.07 |
| 5 | Conscientiousness | 0.06 |
| 6 | Mast. avoidance | 0.04 |
| 7 | Openness | 0.04 |
| 8 | Num. semester | 0.04 |
| 9 | Work to study | 0.03 |
| 10 | Perf. avoidance | 0.02 |

Import. = Importance

## Prediction rules

| Rules | Import. | Coef. |
|---|---|---|
| Conscientiousness <= 0.0649 & HS GPA <= 0.3529 | 0.11 | -0.23 |
| Perf. approach > -0.6106 & PTAwr goal <= 2 | 0.10 | 0.28 |
| PTAwr goal <= 2 & Perf. approach > -0.6241 | 0.09 | 0.26 |
| HS GPA <= -0.266 & Mast. avoidance <= 0.1248 | 0.08 | -0.20 |
| Openness <= 0.0731 & Non-major <= 0 | 0.08 | 0.15 |
| HS GPA > -0.6857 & Num. semester <= 1 | 0.07 | 0.15 |
| Work to study <= 0 & Non-major <= 0 | 0.06 | 0.11 |
| Perf. approach > -0.6397 & PTAwr goal <= 2 | 0.05 | 0.14 |
| HS GPA <= -0.2982 & Perf. avoidance <= 0.1731 | 0.04 | -0.10 |
| HS GPA <= -0.266 & Mast. avoidance <= 0.5833 | 0.00 | 0.00 |

PTA w.i. = Practice test attempts With incentive

# RQ3 – What predicts the practice test attempts goal status?

What predicts the practice test attempts goal?

**Figure 7:** Interaction between HS GPA & Cons.

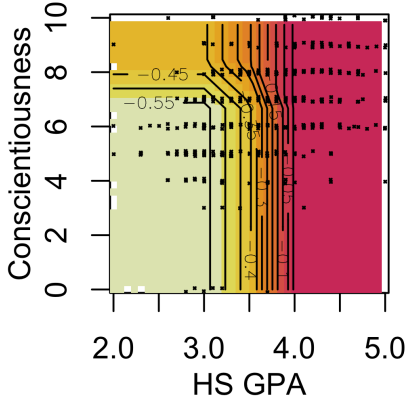
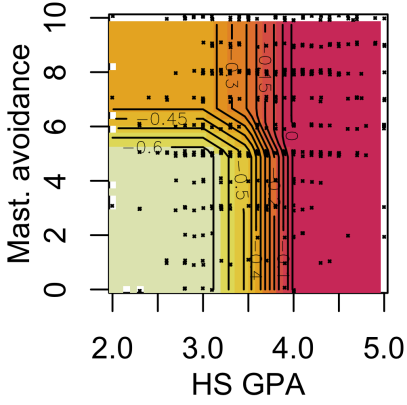



**Figure 8:** Interaction between Mastery Avoidance & HS GPA



## RQ3 – What predicts the practice test attempts goal status?

**Figure 9:** Interaction between Mastery approach & HS GPA

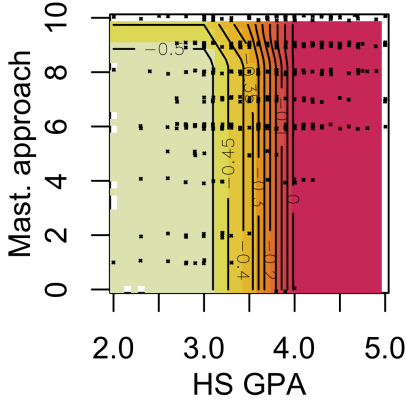



**Figure 10:** Interaction between Num. of sem. & HS GPA