Time-economical assessment of intelligence as predictor variable for academic learning: *Is speeded reasoning enough?* [hint: no!]

Peter Edelsbrunner
ETH Zurich

Test of Relational Reasoning (TORR)

Alexander, Dumas et al., 2015; Goldwater & Schalk, 2016; Holyoak & Lu, 2021; Jastrzębski et al., 2023; Kalra & Richland, 2022

Analogy (10 min)

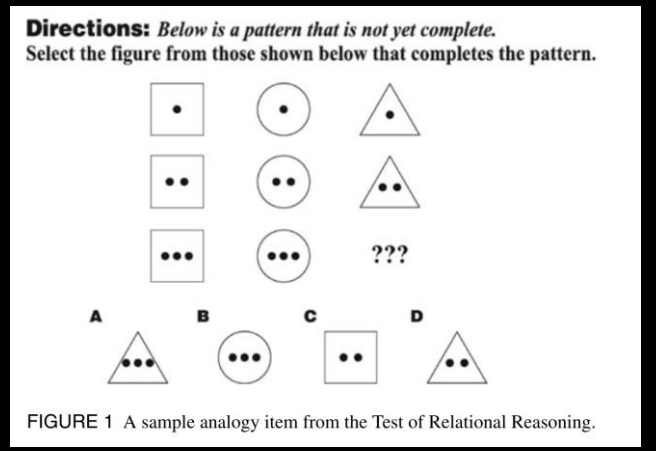


FIGURE 1 A sample analogy item from the Test of Relational Reasoning.

Anomaly (10 min)

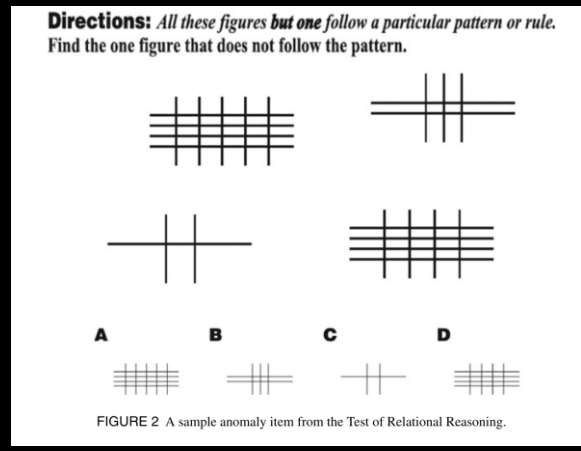
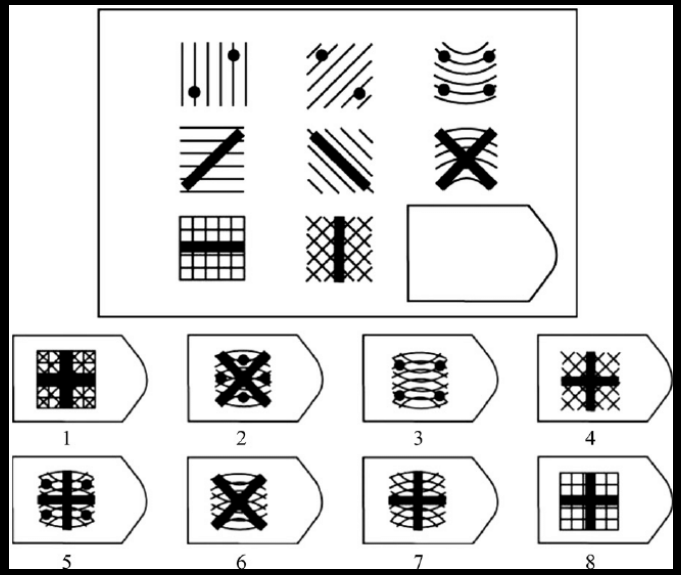


FIGURE 2 A sample anomaly item from the Test of Relational Reasoning.

Raven APM (**40 min; 20 min; 7 min**)

Hamel & Schmittmann, 2006; Poulton et al., 2022



KFT number series (**10 min**)

Heller & Perleth, 2000

7 14 9 18 13 26 21 ?

mini-q
Baudson & Preckel (2015)

64 items
3 minutes
Speeded reasoning

	TRUE	FALSE
The triangle rejects the square	x	
The circle prefers the triangle	x	
The square does not reject the triangle		x

There are no rules to be learned

Speeded reasoning (Baudson & Preckel, 2016)

Analogical mapping of figural and verbal relation ← But it's really trivial!

Does the figural relation match the verbal relation? ← This is also really trivial!

Intelligence 102 (2024) 101804

Contents lists available at ScienceDirect

Intelligence

journal homepage: www.elsevier.com/locate/intell

Don't waste your time measuring intelligence: Further evidence for the validity of a three-minute speeded reasoning test

Anna-Lena Schubert [a,*], Christoph Löffler [a,b], Clara Wiebel [a], Florian Kaulhausen [c], Tanja Gabriele Baudson [d,e]

[a] Department of Psychology, University of Mainz, Mainz, Germany
[b] Department of Psychology, Heidelberg University, Heidelberg, Germany
[c] Department of Psychology, Free University of Berlin, Berlin, Germany
[d] Department of of Human Sciences, Vinzenz Pallotti University, Vallendar, Germany
[e] Institute for Globally Distributed Open Research and Education (IGDORE), Sweden

There are no rules to be learned

Speeded reasoning (Baudson & Preckel, 2016)

Analogical mapping of figural and verbal relation

Does the figural relation match the verbal relation?

The triangle rejects the square
The circle prefers the triangle
The square does not reject the triangle

2 samples, N = 140/505, Age 18-60

«Excellent reliability» (.96-.99)

«Substantial correlation with general cognitive ablities» (r = .57)

«Working memory capacity accounted for the majority (54%) of the association between test performance and general cognitive abilities, whereas individual differences in processing speed did not contribute to this relationship»

«Supporting potential as brief assessment/screening of general cognitive abilities»

«Brief, reliable, valid measure of GCA»

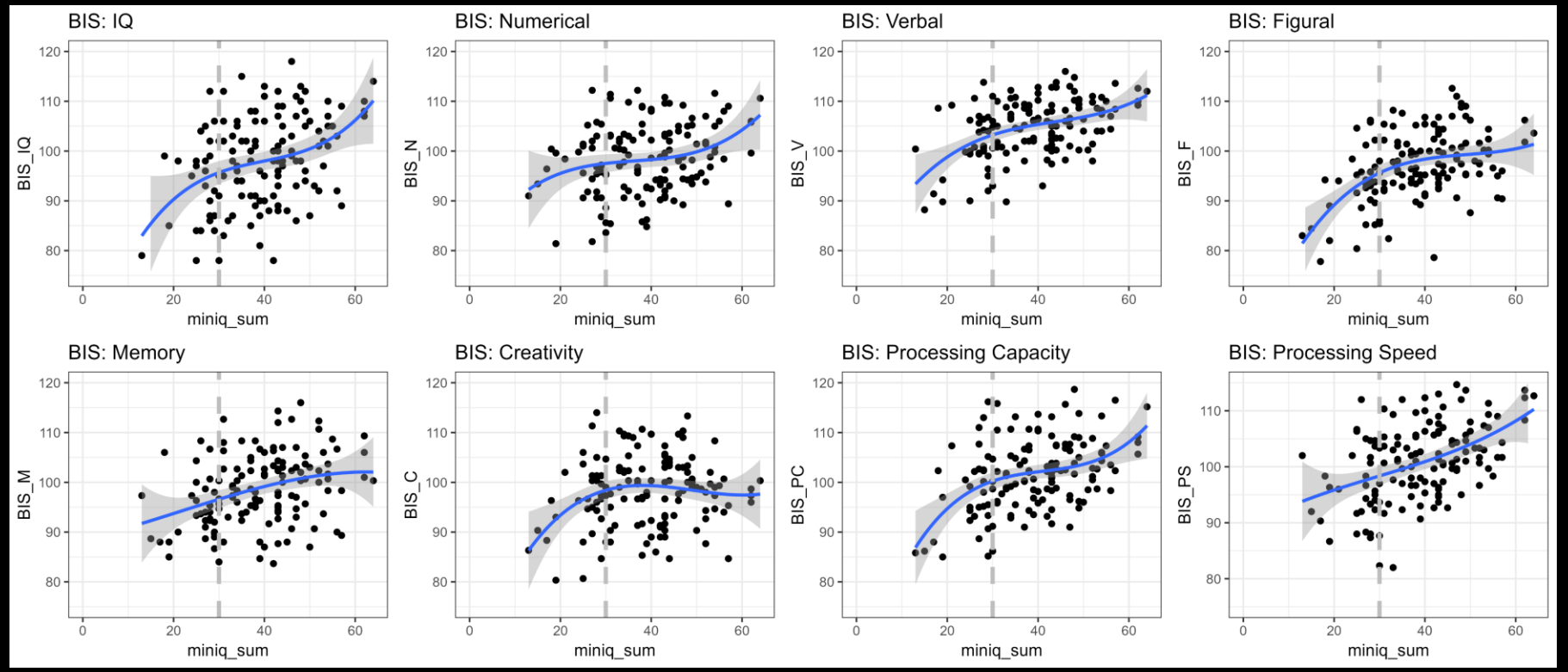
What is the mini-q used for?

Measure of fluid intelligence (Pollet & Schnell, 2017)
Measure of mental abilities (Baltes & Diehl, 2018)
Measure for cognitive abilities (Meuer & Imhoff, 2021)

Relations with broad intelligence measures: Berlin Intelligence Structure Test (BIS)

Beauducel & Kersting, 2002

r = .57 *r* = .26 *r* = .45 *r* = .41



r = .33 *r* = .12 *r* = .41 *r* = .46

- Stronger **figural** and **verbal** than numerical components
- There are **quadratic relations** with most intelligence facets (but speed)
- Relations in the **upper range** are **less pronounced**

Validity: Correlations of the mini-q
Evidence from prior studies

Convergent:

Intelligence structure test: *r* = .25/.47 (Weise et al., 2024) <- *less speeded measure*

Complex problem solving: *r* = .18/.28 (Gigl et al., n.d.) <- *power measure*

GPA: *r* = .07/.25 [parallel version] (Weise et al., 2024) <- *eat this! (Meta-analytic IQ/SA: r = .54; Roth et al., 2015)*

Age: *r* = -.29/-.45/-.50 (e.g., Schubert et al., 2023)

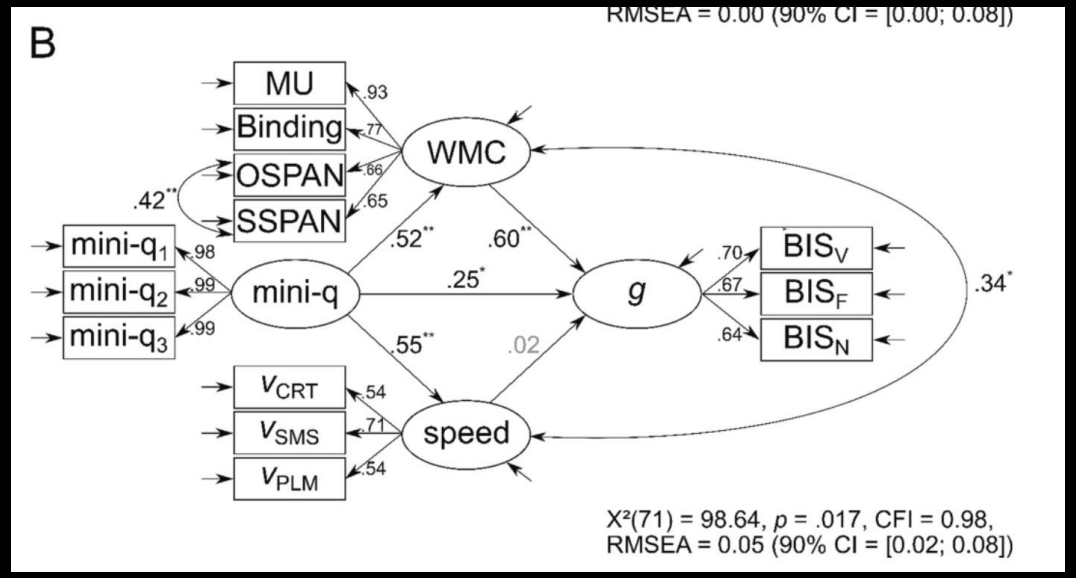
Memory updating: *r* = .47 (Schubert et al., 2023) <- *interesting! Makes sense – so many items in so little time!*

Reaction time measures (working memory; Posner): *r* = -.35 - -.45 (Schubert et al., 2023) <- *higher than with accuracy!*

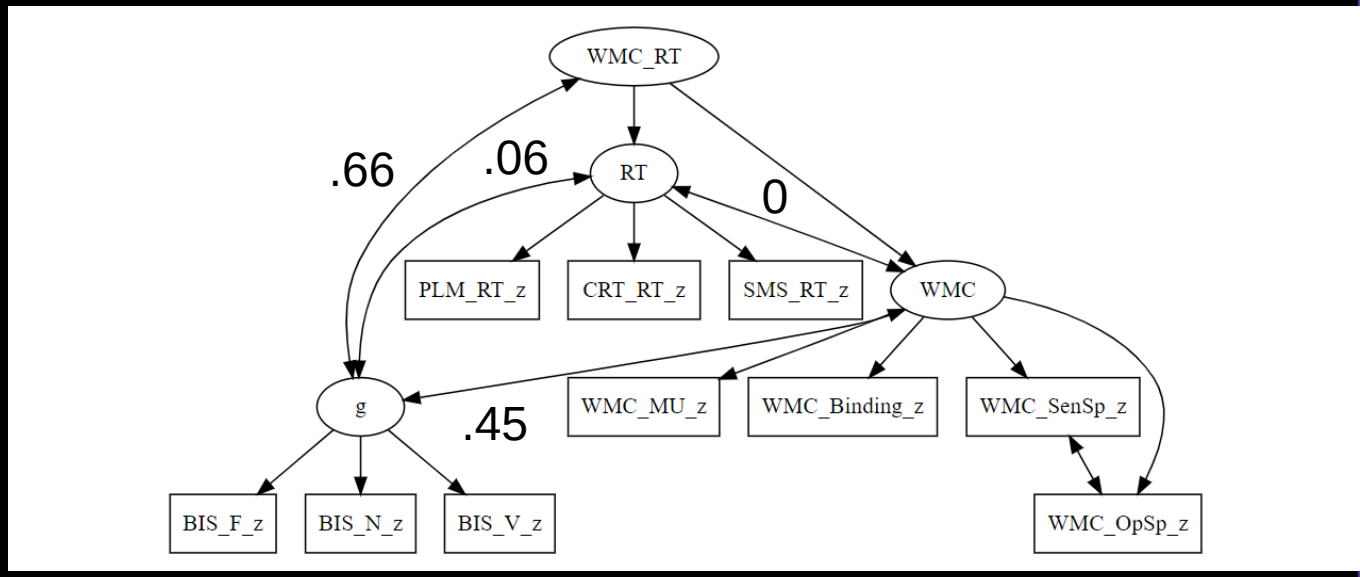
The triangle rejects the square
The circle prefers the triangle
The square does not reject the triangle

The mini-q measures a lot of **processing speed**, **updating, attention**, and rather **little G**, **complex problem solving**, or **GPA-related abilities**

*«**Working memory capacity accounted for** the majority (54%) of the **association between test performance and general cognitive abilities**, whereas individual differences in processing **speed did not contribute** to this relationship»*



Re-analysis with common WMC/RT-factor: «Working memory capacity and processing speed both together accounted for 44% for variation in cognitive abilities and had similar relations with the mini-q



Reliability of the mini-q
Evidence from prior studies

Odd/even: r = .96/.99 (Weise et al., 2024)

Alpha = .94/.96 (Weise et al., 2024)

Alpha = .94/.95 (Wagner & Monzel, 2023)

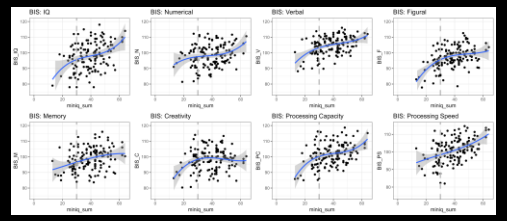
For speeded tests, estimates of internal consistency are invalid and inflated (Anastasi & Drake, 1954)

Retest-reliability should be used for reliability estimation

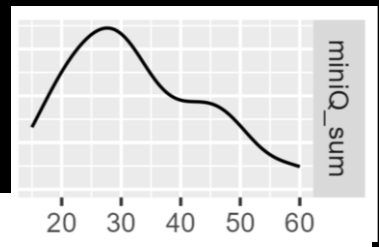
Parallel test forms reliability: r = .76/.80 (Weise et al., 2024)

Relations with broad intelligence measures: Berlin Intelligence Structure Test (BIS)

Beauducel & Kersting, 2002



mini-q distributions in Schubert et al. (2023)



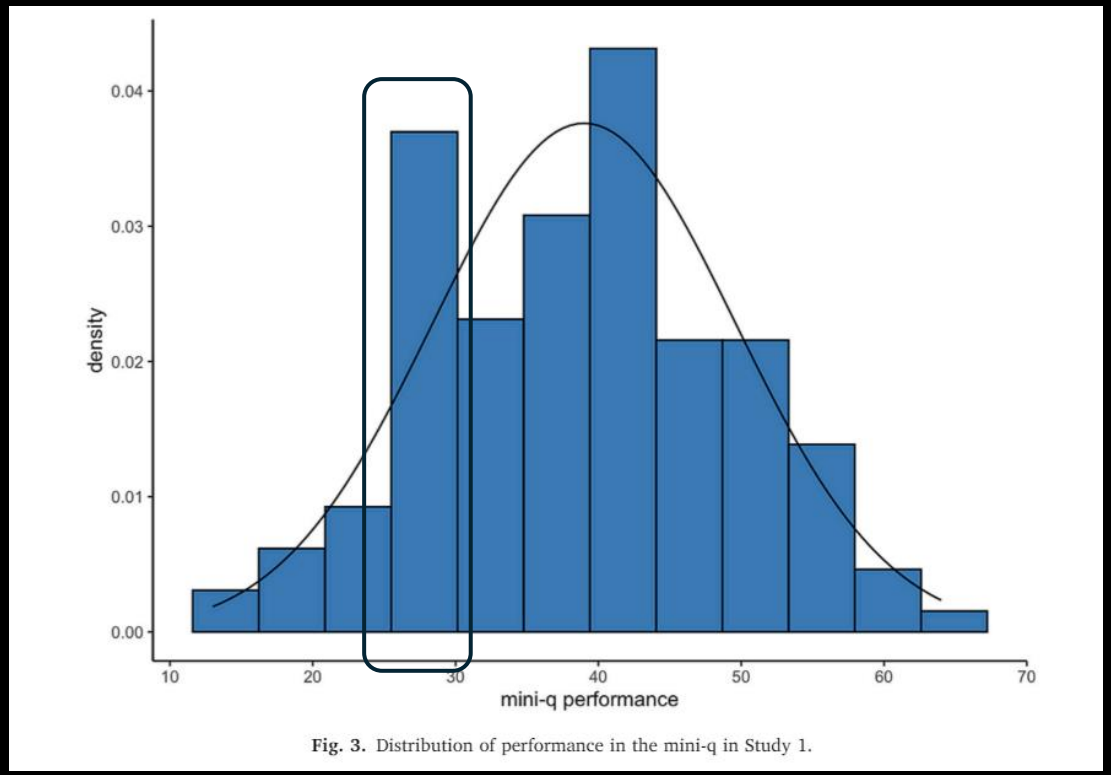


Fig. 3. Distribution of performance in the mini-q in Study 1.

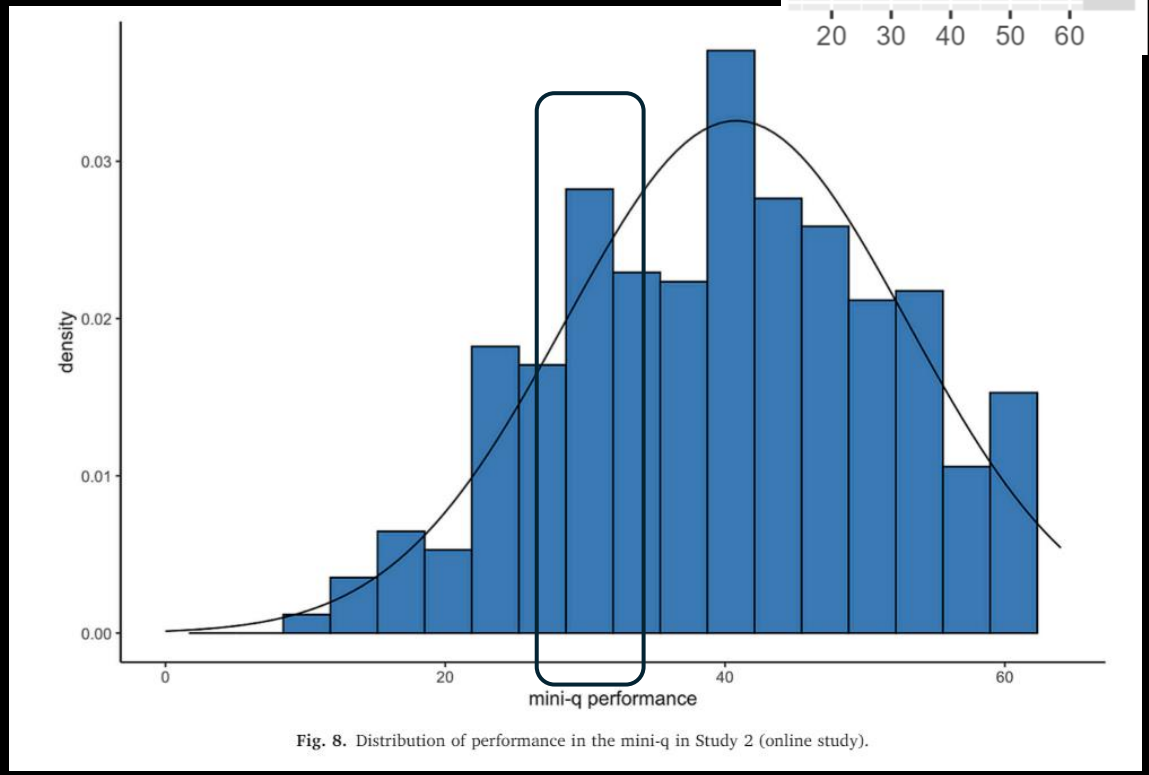


Fig. 8. Distribution of performance in the mini-q in Study 2 (online study).

Bi-/multimodality: Breakdown after speed is the only factor? Item wording effects?

Different strategies? *«I first looked at the figures, but for the later items I started with the sentences»*

An alternative proposal: Test of Relational Reasoning (TORR; Alexander, Dumas et al., 2015)

Analogy

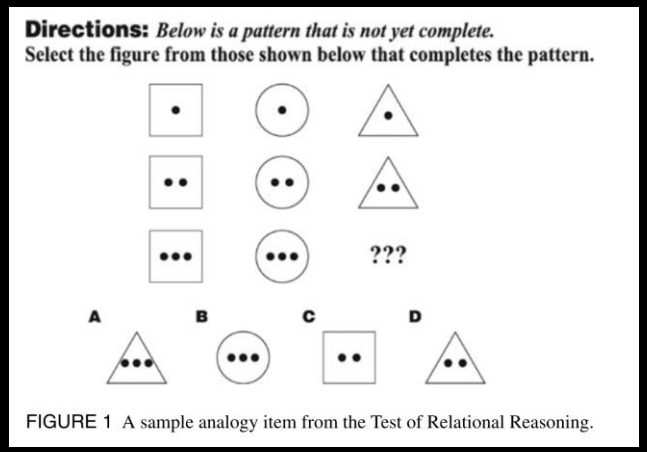


FIGURE 1 A sample analogy item from the Test of Relational Reasoning.

Anomaly

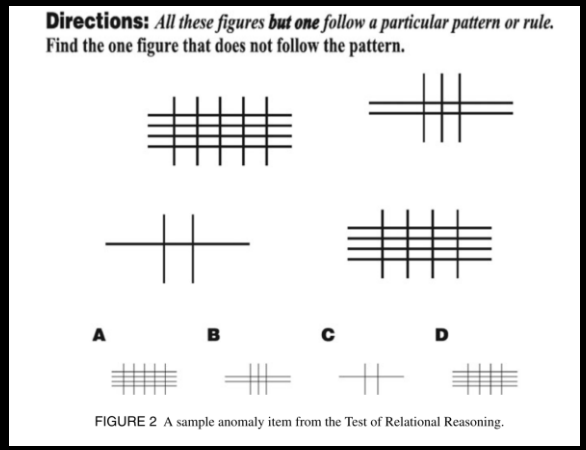


FIGURE 2 A sample anomaly item from the Test of Relational Reasoning.

N = 98 high school students

Representational competence

Conceptual Knowledge

Mastery goal orientation

	Anal_mean	Anom_mean	miniQ_sum
RCFI	Corr: 0.459*	Corr: 0.424*	Corr: 0.047
KWTPre	Corr: 0.264	Corr: 0.281	Corr: 0.255
KWTPost	Corr: 0.223	Corr: 0.298	Corr: 0.189
Note	Corr: -0.070	Corr: 0.240	Corr: 0.037
Model_mean	Corr: 0.201	Corr: -0.070	Corr: -0.080
Mastery_mean	Corr: 0.147	Corr: 0.028	Corr: 0.389*

It might be a worse indicator of academic learning than other short reasoning scales

It might draw more upon current achievement motivation (Freund et al., 2015)

Conclusions

- The mini-q correlates with intelligence only in the lower range; it is not a proxy of intelligence

- The mini-q measures mostly processing speed, maybe attention, updating

- Its reliability is maybe around .80 but yet unknown – retest-studies are required

Hypotheses for future research

The mini-q is multidimensional

The mini-q draws on different strategies

Selecting good strategies might be the only intelligent component

The mini-q draws on inhibition (inverse items) /updating

The mini-q draws on current achievement motivation

It might not be the right tool to get a reliable and valid covariate in predicting complex cognitive states and traits

Not everything that requires **some kind of reasoning** measures **intelligence**

Don’t Waste your Time Using the miniQ: It is not an Effective Proxy of Intelligence or Predictor of Academic Learning

Reaction time tasks: Obviously good correlations

WMC and miniq still correlate .34 even after contr
miniq! And their combined variance (i.e., bifactorg
explains some variation in the miniq!

The multiple regression «no explaantory power»-fa
throughout the paper!

Lob:
Important, well-conducted study.

But I disagree with some interpretations for reaso

First:

Second:

Third:

Scatter plots iwth BIS C and F: Wow! Extremely qu

Short forms always been interest

- Raven etc.
- Or just a numeric scale (e.g., number series
- Or just a spatial (rotation) or verbal (analogies) scale, or cubes or smth for fluid stuff mostly
- Alternative approach/idea: Baddeley.
- Mini-q:
- Show 2 quick examples
- What does this measure and what is it good for?
- Overview of available studies: First studies did this and that.
- Recent study: First bigger investigation into… [summarize]
- Let's see:
- The aim is:
- Control for intelligence
- Use it as a predictor
- Use it as a moderator variable? (diff. from predictor?)

- «Screening»? No definition given what this should mean and be good for? 2-Stage: this is the first stage. Or should just be a proxy in a one-stage screening interpretation?

- BIS strongly timed (i.e., speed-component laden) measure

Reaction time tasks: Obviously good correlations

WMC and miniq still correlate .34 even after contr
miniq! And their combined variance (i.e., bifactorg
explains some variation in the miniq!

The multiple regression «no explaantory power»-fa
throughout the paper!

Lob:
Important, well-conducted study.

But I disagree with some interpretations for reaso

First:

Second:

Third:

Scatter plots iwth BIS C and F: Wow! Extremely qu

Correlations of the mini-q
Evidence from prior studies

Divergent (Bertrams et al., 2024):

Social motivation: .06
Self-deceptive enhancement: .02
Mentalising: .02
Perspective taking: .02
Empathic concern: -.03
Communication and reciprocity: .11

Victim sensitivity: .20 (Meuer & Imhoff, 2021)

Moderately negative with (Pollet & Schnell, 2017):
Age, subjective well-being, religiosity, spirituality, practicality, harmony, self-compassion, demands and appreciation in school experience, meaningful work, joy of working

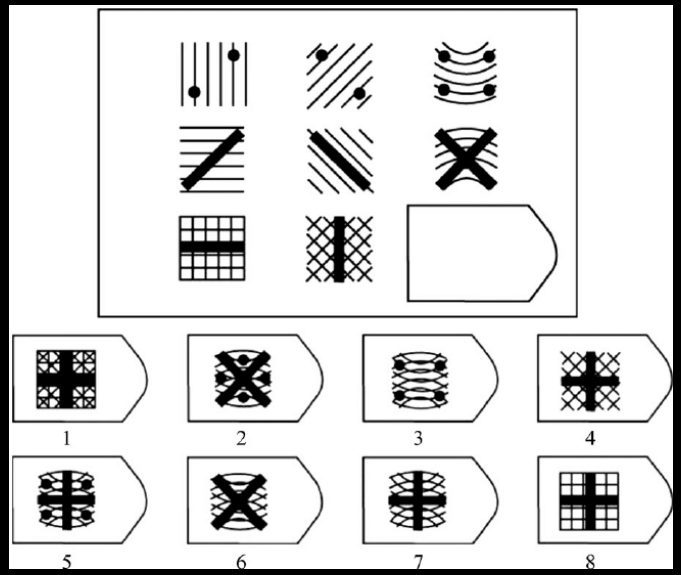
The triangle rejects the square
The circle prefers the triangle
The square does not reject the triangle

Don’t Waste your Time Using the mini-q: It is not an Effective Proxy of Intelligence or Predictor of Academic Learning

Peter Edelsbrunner
ETH Zurich

Raven APM (**40 min; 20 min; 7 min**)

Hamel & Schmittmann, 2006; Poulton et al., 2022



KFT number series (**10 min**)

Heller & Perleth, 2000

7 14 9 18 13 26 21 ?

mini-q
Baudson & Preckel (2015)

64 items
3 minutes
Speeded reasoning

	TRUE	FALSE
The triangle rejects the square	x	
The circle prefers the triangle	x	
The square does not reject the triangle		x

There are no rules to be learned

Speeded reasoning (Baudson & Preckel, 2016)

Analogical mapping of figural and verbal relation ← But it's really trivial!

Reasoning: Does the figural relation match the verbal relation? ← This is also really trivial!

Intelligence 102 (2024) 101804

Contents lists available at ScienceDirect

Intelligence

journal homepage: www.elsevier.com/locate/intell

Don't waste your time measuring intelligence: Further evidence for the validity of a three-minute speeded reasoning test

Anna-Lena Schubert [a,*], Christoph Löffler [a,b], Clara Wiebel [a], Florian Kaulhausen [c], Tanja Gabriele Baudson [d,e]

[a] Department of Psychology, University of Mainz, Mainz, Germany
[b] Department of Psychology, Heidelberg University, Heidelberg, Germany
[c] Department of Psychology, Free University of Berlin, Berlin, Germany
[d] Department of of Human Sciences, Vinzenz Pallotti University, Vallendar, Germany
[e] Institute for Globally Distributed Open Research and Education (IGDORE), Sweden

There are no rules to be learned

Speeded reasoning (Baudson & Preckel, 2016)

Analogical mapping of figural and verbal relation

Does the figural relation match the verbal relation?

The triangle rejects the square
The circle prefers the triangle
The square does not reject the triangle

2 samples, N = 140/505, Age 18-60

«Excellent reliability» (.96-.99)

«Substantial correlation with general cognitive ablities» (r = .57)

«Working memory capacity accounted for the majority (54%) of the association between test performance and general cognitive abilities, whereas individual differences in processing speed did not contribute to this relationship»

«Supporting potential as brief assessment/screening of general cognitive abilities»

«Brief, reliable, valid measure of GCA»

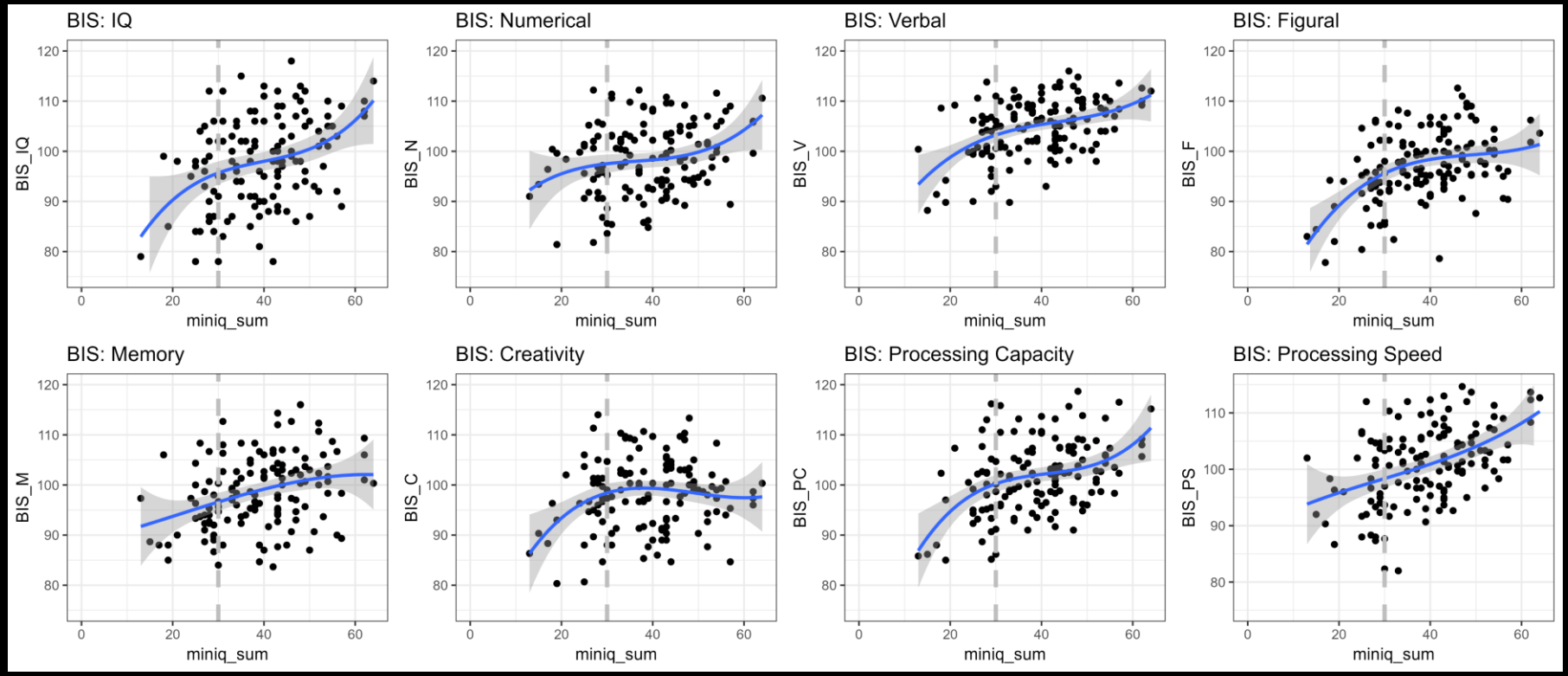
What is the mini-q used for?

Measure of fluid intelligence (Pollet & Schnell, 2017)
Measure of mental abilities (Baltes & Diehl, 2018)
Measure for cognitive abilities (Meuer & Imhoff, 2021)

Relations with broad intelligence measures: Berlin Intelligence Structure Test (BIS)

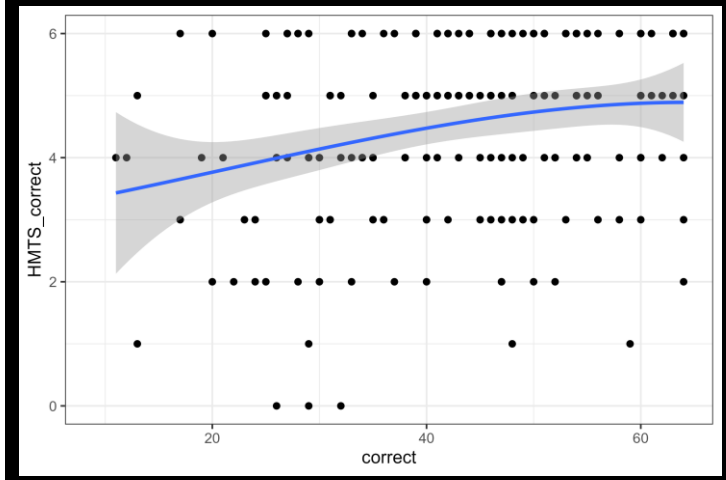
Beauducel & Kersting, 2002

r = .57 *r* = .26 *r* = .45 *r* = .41



r = .33 *r* = .12 *r* = .41 *r* = .46

- Stronger figural and verbal than numerical components
- There are quadratic relations with most intelligence facets
- Relations in the upper range are less pronounced



Relation with short-form of Hagen matrices test (six items)

Schubert et al. 2023

Correlations of the mini-q
Evidence from prior studies

Convergent:

Intelligence structure test: *r* = .25/.47 (Weise et al., 2024) <- *less speeded measure*

Complex problem solving: *r* = .18/.28 (Gigl et al., n.d.) <- *power measure*

GPA: *r* = .07/.25 (Weise et al., 2024) <- *eat this!*

Age: *r* = -.29/-.45/-.50 (e.g., Schubert et al., 2023)

Memory updating: *r* = .47 (Schubert et al., 2023) <- *interesting! Makes sense – so many items in so little time!*

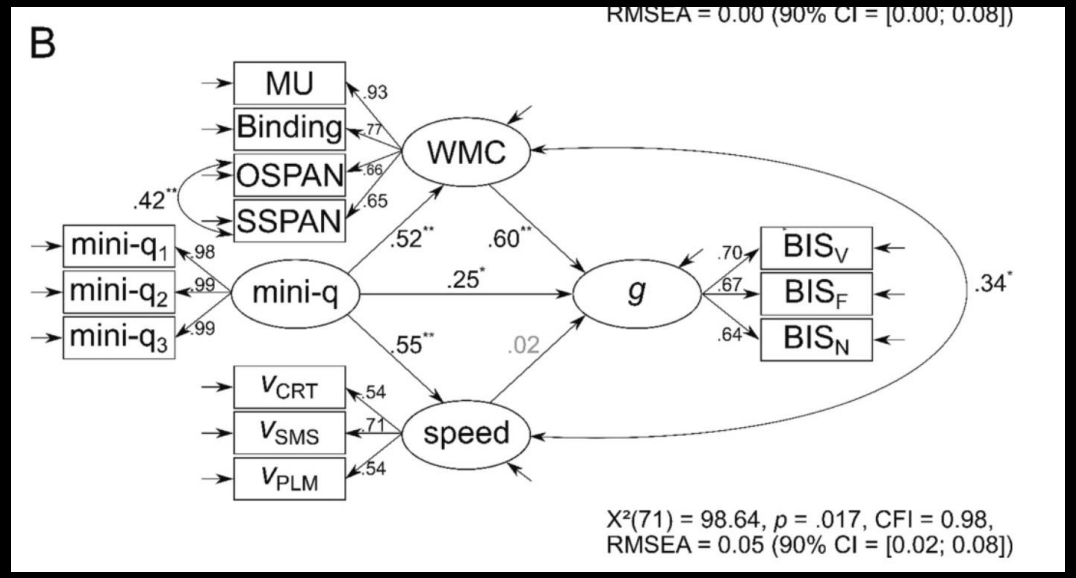
Reaction time measures (working memory; Posner): *r* = -.35 - -.45 (Schubert et al., 2023) <- *higher than with accuracy!*

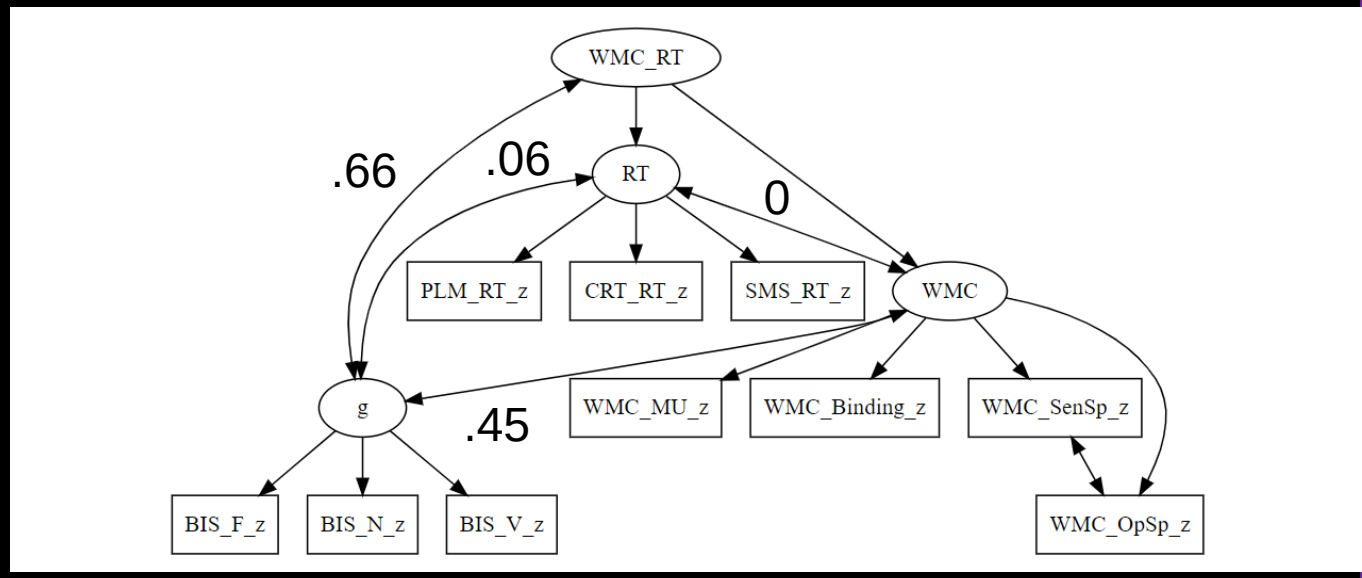
The triangle rejects the square
The circle prefers the triangle
The square does not reject the triangle

The mini-q measures a lot of processing speed, updating, attention, and rather little

«Working memory capacity accounted for the majority (54%) of the association between test performance and general cognitive abilities, whereas individual differences in processing speed did not contribute to this relationship»

«Working memory capacity and processing speed both together accounted for 44% for variation in cognitive abilities and had similar relations with the mini-q





Reliability of the mini-q
Evidence from prior studies

Parallel test forms reliability: r = .76/.80 (Weise et al., 2024)

Odd/even: r = .96/.99 (Weise et al., 2024)

Alpha = .94/.96 (Weise et al., 2024)

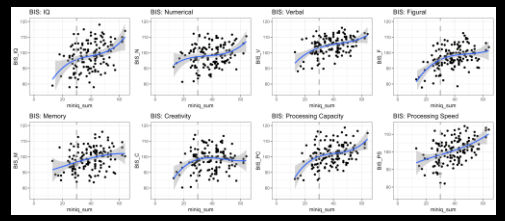
Alpha = *.94/.95* (Wagner & Monzel, 2023)

For speeded tests, estimates of internal consistency are invalid and inflated (Anastasi & Drake, 1954)

Retest-reliability should be used for reliability estimation

Relations with broad intelligence measures: Berlin Intelligence Structure Test (BIS)

Beauducel & Kersting, 2002



mini-q distributions in Schubert et al. (2023)

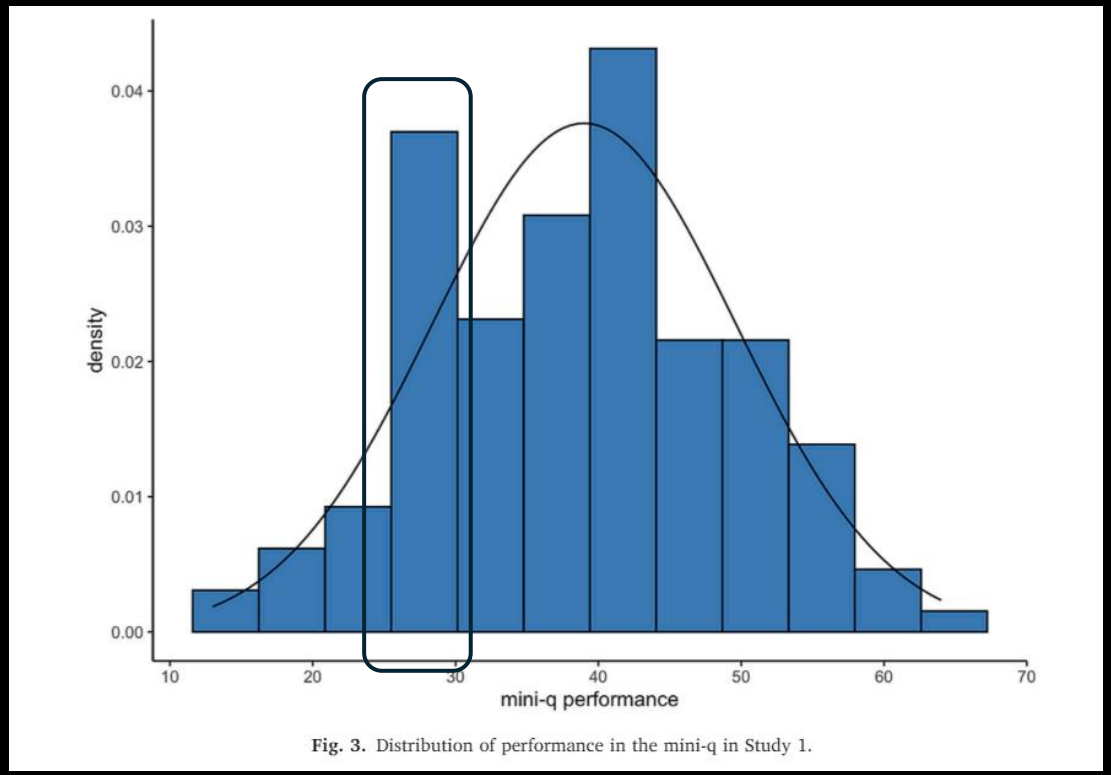


Fig. 3. Distribution of performance in the mini-q in Study 1.

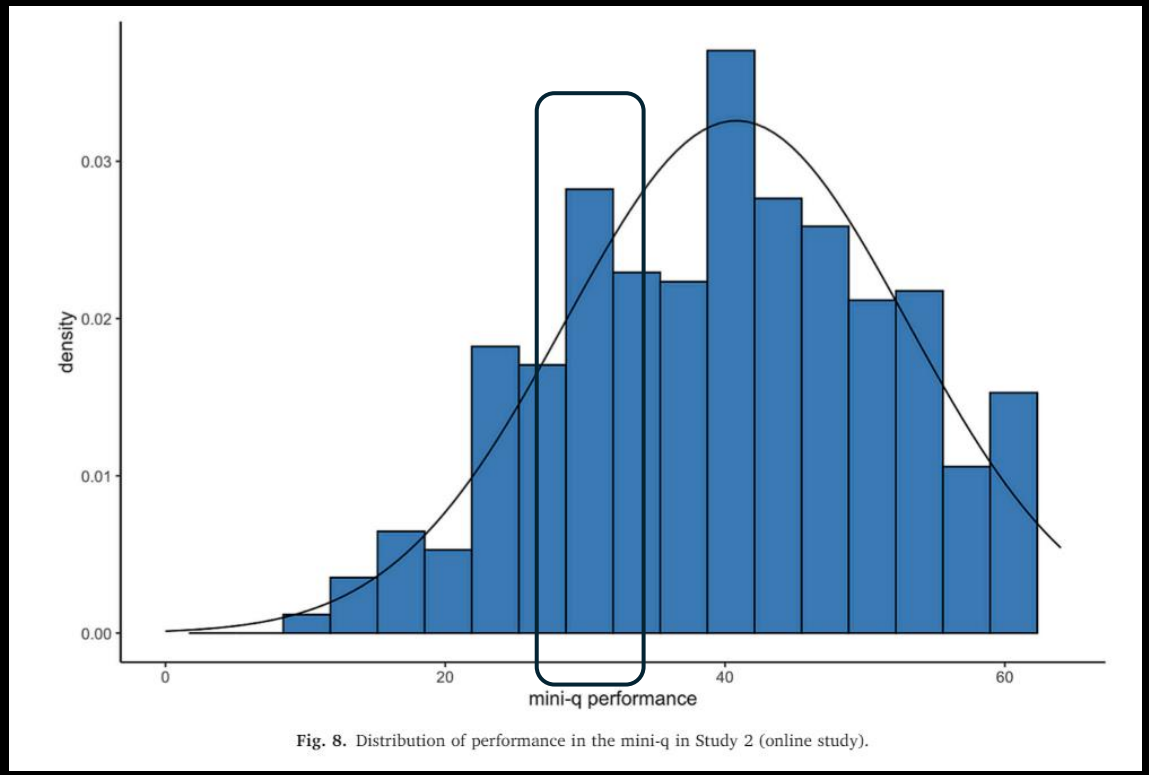


Fig. 8. Distribution of performance in the mini-q in Study 2 (online study).

Bi-/multimodality: Breakdown after speed is the only factor? Item wording effects?

Different strategies? *«I first looked at the figures, but for the later items I started with the sentences»*

An alternative proposal: Test of Relational Reasoning (TORR; Alexander, Dumas et al., 2015)

Analogy

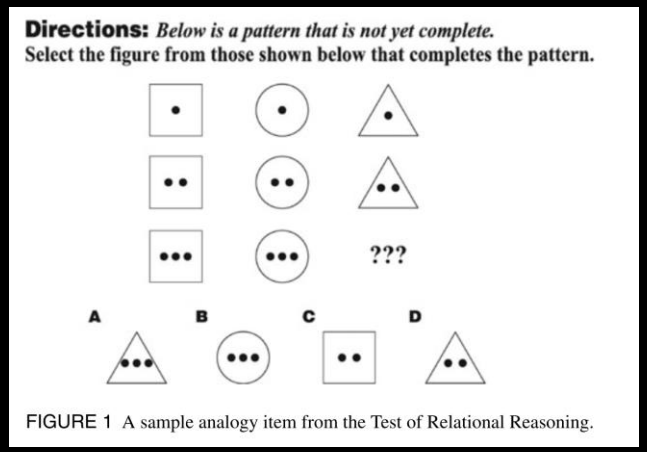


FIGURE 1 A sample analogy item from the Test of Relational Reasoning.

Anomaly

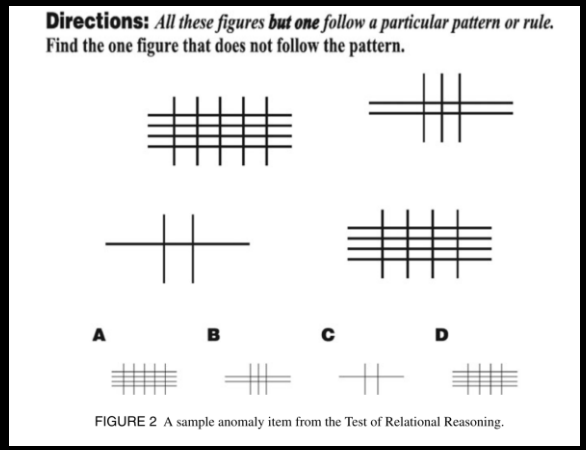


FIGURE 2 A sample anomaly item from the Test of Relational Reasoning.

N = 98 high school students

Representational competence

Conceptual Knowledge

Mastery goal orientation

	Anal_mean	Anom_mean	miniQ_sum
RCFI	Corr: 0.459*	Corr: 0.424*	Corr: 0.047
KWTPre	Corr: 0.264	Corr: 0.281	Corr: 0.255
KWTPost	Corr: 0.223	Corr: 0.298	Corr: 0.189
Note	Corr: -0.070	Corr: 0.240	Corr: 0.037
Model_mean	Corr: 0.201	Corr: -0.070	Corr: -0.080
Mastery_mean	Corr: 0.147	Corr: 0.028	Corr: 0.389*

It might be a worse indicator of academic learning than other short reasoning scales

It might draw more upon current achievement motivation (Freund et al., 2015)

Conclusions

- The mini-q correlates with intelligence only in the lower range; it is not a proxy of intelligence

- The mini-q measures mostly processing speed

- Its reliability is maybe around .80 but yet unknown – retest-studies are required

Hypotheses for future research

The mini-q is multidimensional

The mini-q draws on different strategies

Selecting good strategies might be the only intelligent component

The mini-q draws on inhibition

The mini-q draws on current achievement motivation

It might not be the right tool to get a reliable and valid covariate in predicting complex cognitive states and traits

Not everything that requires **some kind of reasoning** measures **intelligence**

Don’t Waste your Time Using the miniQ: It is not an Effective Proxy of Intelligence or Predictor of Academic Learning

Reaction time tasks: Obviously good correlations

WMC and miniq still correlate .34 even after contr
miniq! And their combined variance (i.e., bifactorg
explains some variation in the miniq!

The multiple regression «no explaantory power»-fa
throughout the paper!

Lob:
Important, well-conducted study.

But I disagree with some interpretations for reaso

First:

Second:

Third:

Scatter plots iwth BIS C and F: Wow! Extremely qu

Short forms always been interest

- Raven etc.
- Or just a numeric scale (e.g., number series
- Or just a spatial (rotation) or verbal (analogies) scale, or cubes or smth for fluid stuff mostly
- Alternative approach/idea: Baddeley.
- Mini-q:
- Show 2 quick examples
- What does this measure and what is it good for?
- Overview of available studies: First studies did this and that.
- Recent study: First bigger investigation into… [summarize]
- Let's see:
- The aim is:
- Control for intelligence
- Use it as a predictor
- Use it as a moderator variable? (diff. from predictor?)

- «Screening»? No definition given what this should mean and be good for? 2-Stage: this is the first stage. Or should just be a proxy in a one-stage screening interpretation?

- BIS strongly timed (i.e., speed-component laden) measure

Reaction time tasks: Obviously good correlations

WMC and miniq still correlate .34 even after contr
miniq! And their combined variance (i.e., bifactorg
explains some variation in the miniq!

The multiple regression «no explaantory power»-fa
throughout the paper!

Lob:
Important, well-conducted study.

But I disagree with some interpretations for reaso

First:

Second:

Third:

Scatter plots iwth BIS C and F: Wow! Extremely qu

Correlations of the mini-q
Evidence from prior studies

Divergent (Bertrams et al., 2024):

Social motivation: .06
Self-deceptive enhancement: .02
Mentalising: .02
Perspective taking: .02
Empathic concern: -.03
Communication and reciprocity: .11

Victim sensitivity: .20 (Meuer & Imhoff, 2021)

Moderately negative with (Pollet & Schnell, 2017):
Age, subjective well-being, religiosity, spirituality, practicality, harmony, self-compassion, demands and appreciation in school experience, meaningful work, joy of working

The triangle rejects the square
The circle prefers the triangle
The square does not reject the triangle