What Intelligence Tests Miss:
Individual Differences in Reasoning Beyond IQ

Keith E. Stanovich

Collaborators:

Richard F. West

Maggie T. Toplak

What Intelligence Tests Miss:

The Psychology of Rational Thought:

Yale University Press, 2009

People are confused about the concept of intelligence and about what IQ tests measure.

The case of George W. Bush

"Too many good docs are getting out of the business. Too many OB-GYNs aren't able to practice their love with women all across this country."
George W. Bush — Sept. 6, 2004

“Our enemies are innovative and resourceful, and so are we. They never stop thinking about new ways to harm our country and our people, and neither do we”
George W. Bush, Washington, DC, Aug. 5, 2004

“I know what I believe. I will continue to articulate what I believe--I believe what I believe is right”
George W. Bush, Rome, Italy, July 22, 2001

“Do you have blacks too?”
George Bush to Brazilian President Fernando Cardoso, Washington, DC, Nov. 8, 2001

“I’m also not very analytical. You know I don’t spend a lot of time thinking about myself, about why I do things.”-- President George W. Bush, aboard Air Force One, June, 4, 2003

Both opponents *and* supporters of Bush agree that his thinking is defective in certain ways and both assume that he would not score highly on conventional IQ tests

"he is impatient and quick to anger; sometimes glib, even dogmatic; often uncurious and as a result ill-informed"
(David Frum, p. 272, 2003)

"has neither the inclination nor the ability to make sophisticated judgments about competing approaches to construing the Constitution"
(George Will, p. 23, 2005)

David Kay “left the meeting almost shocked at Bush’s lack of inquisitiveness” (Bob Woodword, 2006, p. 237)

Senior officials in Baghdad noted the President’s “obvious lack of interest in long, detailed discussions, had a chilling effect….by all accounts, he is not intellectually curious. Outsiders brought into the Bush Bubble have observed that faith, not evidence, is the basis for decision making” (Newsweek, Dec. 19, 2005, p. 37).

“when any doubt started to creep into the small, windowless Situation Room, the president almost stomped it out” (p. 371, Bob Woodward, 2006).

lack of intellectual engagement

cognitive inflexibility

high need for closure

high belief perseverance

overconfidence

insensitivity to inconsistency

Pro-rated scores from college entrance exams and Armed Forces placement tests show:

Al Gore IQ: 125-135

John Kerry IQ: 120-125

George W. Bush IQ: 120-125

Everyone, both opponents *and* supporters, got it wrong about Bush's intelligence.

But for totally different reasons. Both opponents and supporters of Bush are confused about what intelligence tests measure but they are confused in different ways.

Bush's Detractors

Thus, the Bush detractors must have assumed that a mental quality (rational thinking tendencies) could be detected by the tests that in fact the tests do not detect at all.

In contrast, Bush’s supporters like his actions but admit that he has “street smarts,” or common sense, rather than “school smarts.”

That the tests would actually measure a quality that cast Bush in a favorable light was something his supporters never anticipated.

Both groups--supporters and opponents of Bush--are confused about what intelligence tests do and do not measure.

Both have become confused by what I have termed the greatest historical anomaly in modern psychological science.

The Anomaly:
The 2002 Nobel Prize in Economics

"the analysis of human judgment and decision-making by cognitive psychologists" (The Royal Swedish Academy of Sciences, 2002).

One reason that this work was so influential was that it spoke to deep issues concerning human rationality.

As the Nobel announcement noted, “Kahneman and Tversky discovered how judgment under uncertainty systematically departs from the kind of rationality postulated in traditional economic theory” (The Royal Swedish Academy of Sciences, 2002).

Errors of Judgment and Decision Making

It has been shown that there are systematic differences among individuals in the tendency to make errors of judgment and decision making. My own research group has tried to find out what predicts these individual differences (Stanovich & West, 1997, 1998, 2000, 2008; West, Toplak, & Stanovich, 2009).

That there are systematic individual differences in the judgment and decision making situations studied by Kahneman and Tversky means that there are variations in important attributes of human cognition related to rationality. It is a curious fact that none of these critical attributes of human thinking are assessed on IQ tests (or their proxies such as the SAT).

“Good Thinking” and Judgment and Decision Making

In fact, the type of “good thinking” that Kahneman and Tversky studied was deemed so important that research on it was awarded the Nobel Prize.

Yet assessments of such good thinking are nowhere to be found on IQ tests.

It is a profound historical irony of the behavioral sciences that the Nobel Prize was awarded for studies of cognitive characteristics that are entirely missing from the most well-known mental assessment device in the behavioral sciences—the intelligence test.

Because of their vast influence, IQ tests have both explicitly and implicitly defined, for the layperson and psychologist alike, what cognitive attributes to value. This is the social context of our conclusion that….

Intelligence tests are radically incomplete as measures of cognitive functioning.

At this point, someone will usually ask:

So what? Doesn’t everyone know that IQ scores are useless and that they don’t predict much?

Didn’t Howard Gardner and other critics of IQ tests already show that they are pretty useless?

Doesn’t everybody know this already?

Well,

What everyone knows is wrong.

IQ tests predict performance on an almost limitless number of cognitive tasks.

Intelligence Tests Predict Performance on an Almost Limitlesss Number of Cognitive Tasks

Sampling of Cognitive Tasks that Correlate with Intelligence (from Carroll, 1993)

General Sequential Reasoning	Associative Memory	Speech sound Discrimination	Ideational Fluency
Induction	Free Recall Memory	General Sound Discrimination	Naming Fluency
Quantitative Reasoning	Meaningful Memory	Sound-Frequency Discrimination	Associational Fluency
Piagetian Reasoning	Visual Memory	Sound-Intensity/Duration Discrim.	Expressional Fluency
Language Development	Learning Ability	Musical Discrimination & Judgment	Word Fluency
Language Comprehension	Spatial Relations	Resistance to Auditory Stimulus Distortion	Sensitivity to Problems
Lexical Knowledge	Closure Speed	Temporal Tracking	Figural Fluency
Reading Comprehension	Flexibility of Closure	Maintaining & Judging Rhythm	Figural Flexibility
Reading Decoding	Serial Perceptual Integration	Memory for Sound Patterns	Rate of Test Taking
Cloze Ability	Spatial Scanning	Absolute Pitch	Numerical Facility
Spelling Ability	Perceptual Speed	Sound Localization	Simple Reaction Time
Phonetic Coding	Imagery		Choice reaction Time
Grammatical Sensitivity	Length Estimation		Semantic Processing Speed
Foreign Language Aptitude	Perception of Illusions		Mental Comparison Speed
Communication Ability	Perceptual Alternations		
Listening Ability			
Foreign Language Proficiency			
Reading Speed			
Oral Production and Fluency			
Writing Ability			

IQ tests predict not just cognitive tasks in the laboratory. They are the single best predictor of performance in many “real world” occupational settings.

This article presents the research evidence that GMA predicts both occupational level attained and performance within one's chosen occupation and does so better than any other ability, trait, or disposition and better than job experience. The sizes of these relationships with GMA are also larger than most found in psychological research.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162-173.

Table 2

Validity of the General Mental Ability (GMA) Measure in the General Aptitude Test Battery

Complexity level of job[a]	% of workforce	Performance measures	
		On the job	In training
1	14.7	.58	.59
2	2.5	.56	.65
3	62.7	.51	.57
4	17.7	.40	.54
5	2.4	.23	NR

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162-173.

Table 3

The Relation Between General Mental Ability (GMA) and Performance in Job Training and on the Job: Representative Findings From Meta-Analyses

Study	Occupation	Performance measures	
		On the job	In training
Hunter and Hunter (1984)	Medium complexity[a]	.51	.57
Pearlman et al. (1980)	Clerical	.52	.71
Hirsh et al. (1986)	Law enforcement	.38	.76
McHenry et al. (1990)	Military—enlisted	.63[b]	NR
McHenry et al. (1990)	Military—enlisted	.65[c]	NR
Hunter (1986)	Military—enlisted	NR	.63
Ree et al. (1994)	Military—enlisted	.45	NR
Ree and Earles (1991)	Military—enlisted	NR	.60
Schmidt et al. (1979)	First-line supervisors	.64	NR
Schmidt et al. (1979)	Administrative clerks	.67	NR
Schmidt et al. (1980)	Computer programmers	.73	NR
Callender and Osburn (1981)	Refinery workers	.31	.50

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162-173.

The Correlation Between General Mental Ability (GMA) and Job Performance Ratings for Job Incumbents With Various Levels of Job Experience

Years of experience	Total sample size	GMA with performance correlation
0–3	4,424	.35
3–6	3,297	.37
6–9	570	.44
9–12	84	.44
12+	22	.59

The correlation between IQ and any cognitive task is so ubiquitous that cognitive and developmental researchers routinely control for IQ when examining any new association involving a cognitive variable.

Spearman's Positive Manifold: 1904

Virtually all mental tests, however diverse they appear on the surface, correlate significantly with one another.
(pp. 212-213)

Mackintosh, N. J. (1998). *IQ and Human Intelligence*.
Oxford: Oxford University Press.

In psychometrics, the concept of general intelligence derives from the observation of ubiquitous positive correlations among different kinds of cognitive tests: **to some extent, people who do well on one test also tend to do well on others** (p. 132)

Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., Cox, S., et al. (2008). Goal neglect and Spearman's g: Competing parts of a complex task. *Journal of Experimental Psychology: General*, *137*, 131-148.

What is newsworthy?

To put it colloquially:
It’s news when things don’t correlate with intelligence

Most critics of IQ tests assume that there will be no further news in the cognitive domain

Hence, the popularity of advocating for “intelligences” outside of the cognitive domain: emotional intelligence, social intelligence, bodily kinesthetic intelligence, etc.

However, these standard critiques of intelligence tests contain the unstated assumption that although intelligence tests miss certain key noncognitive areas, they encompass most of what is important cognitively.

In fact, intelligence, as conventionally measured, may be missing a critical *cognitive* domain—a domain of thinking itself.

A major conclusion from our work:

From an individual differences perspective, there may be more news in the cognitive domain.

The case of Bush, and the Nobel Prize to Kahneman, suggest as much

We do not have to move to the social and emotional domains to find things that IQ tests miss.

The case of Bush: irrational thought and action despite more than adequate intelligence (specifically, Gf)

Dysrationalia

The odd historical contingency of our field:

Assessments of good judgment and decision making—the type of thinking that helps us achieve our goals and that Kahneman received the Nobel Prize for studying—are missing from IQ tests.

What is missing from IQ tests:

Assessments of the tendency to think rationally.

The dissociation between rationality and intelligence: theoretical and empirical

The theoretical point is that rationality and intelligence are conceptually,
two different things.

Broad versus narrow definitions of intelligence

For example, it took real chutzpah for David Wechsler to define intelligence in his book as “the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment” (p. 7, 1958) despite authoring an IQ test with his name on it that measured no such thing!

practical intelligence

creative intelligence

verbal-linguistic intelligence

interpersonal intelligence

bodily-kinesthetic intelligence

etc.

Separable??

Our definition of intelligence is a more standard one.

It is the accepted definition in cognitive neuroscience and in the journals of the field (e.g., *Intelligence*).

It is the cognitive constructs that the tests actually measure: computational power (Gf) and acquired knowledge (Gc)

Why rationality and intelligence are not the same thing conceptually

To think rationally means adopting appropriate goals, taking the appropriate action given one's goals and beliefs, and holding beliefs that are commensurate with available evidence.

Although IQ tests do assess the ability to focus on an immediate goal in the face of distraction, they do not assess at all whether a person has the tendency to develop goals that are *rational* in the first place!

Likewise, IQ tests are good measures of how well a person can hold beliefs in short-term memory and manipulate those beliefs, but they do not assess at all whether a person has the tendency to *form* beliefs rationally when presented with evidence.

And again, similarly, intelligence tests are good measures of how efficiently a person processes information that has been provided, but they do not at all assess whether the person is a *critical assessor* of information as it is gathered in the natural environment.

But perhaps the skills that are measured on IQ relate to those things—relate to rational thinking skills even though such skills are not assessed directly on the tests.

This is where empirical research comes in—a good amount of it from my own research lab.

Instrumental and Epistemic Rationality in Cognitive Science:

1. Maximizing goal fulfillment (the axioms of utility theory)

2. How well beliefs map to the world

We turn to:

A rich tradition in the heuristics and biases literature of cognitive science, which in its study of many effects, tasks, and biases of cognition provides the empirical measures of the multifarious concept of rational thought.

Tasks, Effects, and Biases Studied
in the Heuristics and Biases Tradition

Hindsight Bias	Overconfidence Effect
Baserate Neglect	The Conjunction Fallacy
Nonregressive Predictions	Myside Bias
Covariation Estimation	Pseudodiagnosticity
Belief Bias	Inappropriate Anchoring
Illusory Correlation	Belief Perseverance
Preference Reversals	Outcome Bias
Commission Bias	Failure of Inconsistency Detection
Violation of SEU Axioms	Ignoring Denominator of the Likelihood Ratio
Failure to Generate Alternative Explanations	Unrealistic Optimism

The surprising findings from over a decade's worth of work

(Kokis et al., 2002; Sa et al., 1999; Stanovich & West, 1997, 1998, 1999, 2000, 2007, 2008; Toplak & Stanovich, 2002, 2003; West & Stanovich, 2003; West, Toplak, & Stanovich, 2008)

Tasks that Fail to Show Associations with Cognitive Ability

Noncausal Baserate Usage (Stanovich & West, 1998c, 1999, 2008)
Conjunction Fallacy Between-Subjects (S&W, 2008)
Framing Between-Subjects (S&W, 2008)
Anchoring Effect (S&W, 2008)
Evaluability Less is More Effect (S&W, 2008)
Proportion Dominance Effect (S&W, 2008)
Sunk Cost Effect (S&W, 2008; Parker & Fischhoff, 2005)
Risk/Benefit Confounding (S&W, 2008)
Omission Bias (S&W, 2008)
Perspective Bias (S&W, 2008)
Certainty Effect (S&W, 2008)
WTP/WTA Difference (S&W, 2008)
Myside Bias - Between and Within-S (Stanovich & West, 2007, 2008)
NewcombÕs Problem (Stanovich & West, 1999; Toplak & Stanovich, 2002)

Tasks that Show .10-.30 Associations with Cognitive Ability

Causal Baserate Usage (Stanovich & West, 1998c, 1998d)
Outcome Bias (Stanovich & West, 1998c, 2008)
Framing Within-Subjects (Bruine de Bruin, Parker, & Fischhoff, 2007; Frederick, 2005; Parker & Fischhoff, 2005; Stanovich & West, 1998b, 1999)
Denominator Neglect (S&W, 2008; Kokis et al., 2002)
Probability Matching (S&W, 2008; West & Stanovich, 2003)
Hindsight Bias (Stanovich & West, 1998c)
Ignoring P(D/NH) (Stanovich & West, 1998d, 1999)
Covariation Detection (Stanovich & West, 1998c, 1998d; Sa et al., 1999)
Belief Bias in Syllogistic Reasoning (Stanovich & West, 1998c, 2008)
Overconfidence Effect (Bruine de Bruin, Parker, & Fischhoff, 2007)
Informal Argument Evaluation (Stanovich & West, 1997, 2008)
Four Card Selection Task (Stanovich & West, 1998a, 2008)
EV Maximization in Gambles (Frederick, 2005; Benjamin & Shapiro, 2005)

West, R. F., Toplak, M. E., & Stanovich, K. E. (in press). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. Journal of Educational Psychology.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. Journal of Personality and Social Psychology, 94, 672-695.

Stanovich, K. E. (2009). The psychology of rational thought: What intelligence tests miss. New Haven, CT: Yale University Press.

Thinking dispositions (need for cognition, actively openminded thinking, reflectivity/impulsivity) predict performance on heuristics and biases tasks after intelligence has been partialled out:

Bruine de Bruin, Parker, & Fischhoff (JPSP, 2007)
Kokis et al. (JECP, 2002)
Klaczynski & Lavallee (JECP, 2005)
Parker & Fischhoff (JBDM, 2005)
Sa et al. (JEP, 1999)
Stanovich & West (1997, 1998, 2000, 2007)
Toplak & Stanovich (JEP, 2002)
Toplak et al. (JBDM, 2007)

Thinking Dispositions Examined in Various Studies:

Need for cognition

Actively openminded thinking

Belief identification

Rational experiential inventory

Need for closure

Openness

Master rationality motive

Conscientiousness

Hierarchical Regression Analysis Predicting Rational Thinking Composite Score of Performance on All Seven Tasks

Step	Variable	R	R^2 Change	F Change	Final F	Partial r
1.	SAT	.497	.247	172.94***	125.42***	.439
2.	TDC composite	.588	.098	79.03***	79.03***	.361

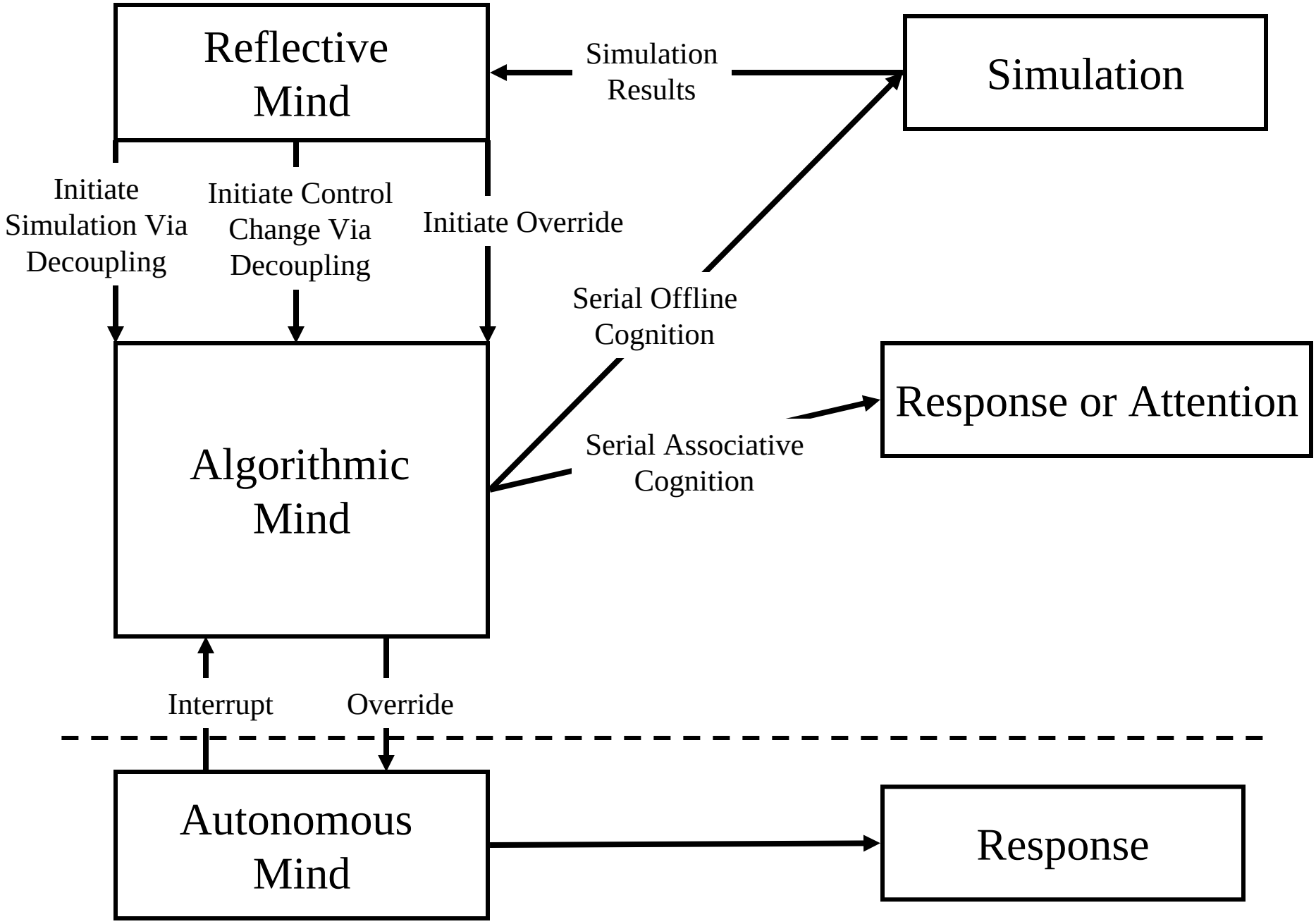
*** $p < .001$

Fig 4. *Simultaneous Regression Analyses on the Inductive, Deductive, and Statistical Task Performance*

	β weight	*t*-value	Unique variance explained	Partial *r*
Criterion Variable = Inductive Problems				
Cognitive Ability	.376	4.02***	.109	.365
AOT Composite	.248	2.65**	.048	.250
Overall Regression: *F* = 21.61***				
Multiple *R* = .540				
Multiple *R*-squared = .292				
Criterion Variable = Deductive Problems				
Cognitive Ability	.313	3.22**	.080	.300
AOT Composite	.251	2.58*	.050	.244
Overall Regression: *F* = 16.18***				
Multiple *R* = .485				
Multiple *R*-squared = .236				
Criterion Variable = Probabilistic Reasoning				
Cognitive Ability	.255	2.79**	.064	.262
Need for Cognition	.227	2.48**	.051	.235
Overall Regression: *F* = 7.65***				
Multiple *R* = .357				
Multiple *R*-squared = .130				
Criterion Variable = Probabilistic Reasoning				
Cognitive Ability	.216	2.24*	.042	.214
Superstitious Thinking	-.199	2.07*	.036	-.198
Overall Regression: *F* = 6.64**				
Multiple *R* = .335				
Multiple *R*-squared = .112				

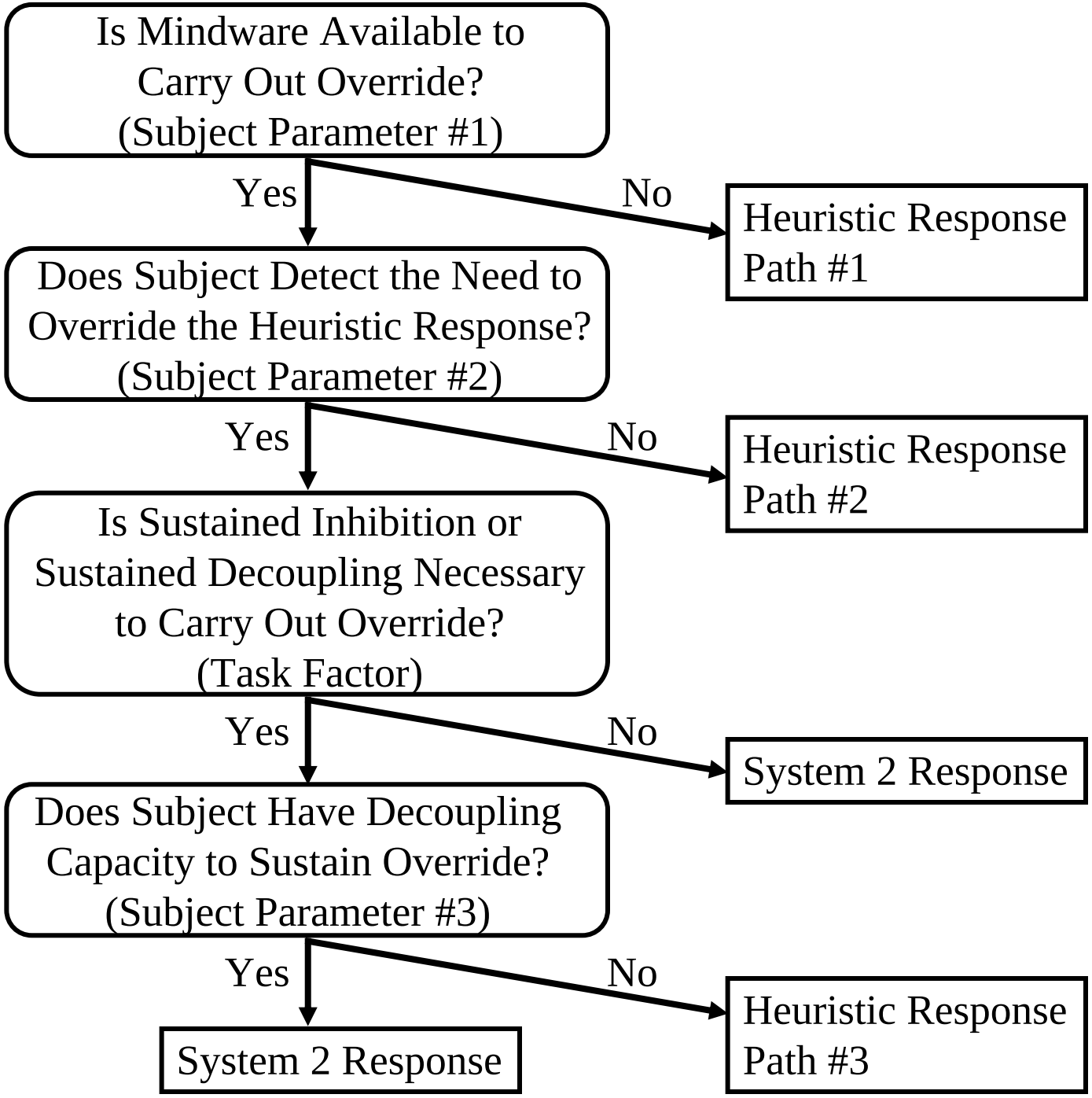
Individual differences in rational thinking and intelligence are not completely overlapping because rationality is a more encompassing concept than intelligence.

Adapt dual-process models of thinking to accommodate our empirical results on patterns of individual differences.


Reflective Mind
Simulation Results
Simulation
Initiate Simulation Via Decoupling
Initiate Control Change Via Decoupling
Initiate Override
Serial Offline Cognition
Response or Attention
Algorithmic Mind
Serial Associative Cognition
Interrupt
Override
Autonomous Mind
Response

Tasks in these Experiments and in Other Studies that do and do not Show Associations with Cognitive Ability

Tasks/Effects that Fail to Correlate with Cognitive Ability	Tasks/Effects that Correlate with Cognitive Ability
Noncausal Baserate Usage (Experiment 1; Stanovich & West, 1998c, 1999)	Causal Baserate Usage (Stanovich & West, 1998c, 1998d)
Conjunction Fallacy Between-Subjects (Experiment 1)	Outcome Bias (Experiment 1; Stanovich & West, 1998c)
Framing Between-Subjects (Experiment 1)	Framing Within-Subjects (Frederick, 2005; Parker & Fischhoff, 2005; Stanovich & West, 1998b, 1999)
Anchoring Effect (Experiment 1)	Denominator Neglect (Experiment 8; Kokis et al., 2002)
Evaluability Less is More Effect (Experiment 2)	Probability Matching (Experiment 8; West & Stanovich, 2003)
Proportion Dominance Effect (Experiment 2)	Hindsight Bias (Stanovich & West, 1998c)
Sunk Cost Effect (Experiment 3; Parker & Fischhoff, 2005)	Ignoring P(D/NH) (Stanovich & West, 1998d, 1999)
Risk/Benefit Confounding (Experiment 4)	Covariation Detection (Stanovich & West, 1998c, 1998d; Sa et al., 1999)
Omission Bias (Experiment 5)	Belief Bias in Syllogistic Reasoning (Experiment 8; Stanovich & West, 1998c)
Perspective Bias (Experiment 5)	Belief Bias in Modus Ponens (Experiment 8)
Certainty Effect (Experiment 6)	Informal Argument Evaluation (Experiment 8; Stanovich & West, 1997)
WTP/WTA Difference (Experiment 6)	Four Card Selection Task (Experiment 8; Stanovich & West, 1998a)
Myside Bias - Between and Within-S (Experiment 7; Stanovich & West, 2007)	EV Maximization in Gambles (Frederick, 2005; Benjamin & Shapiro, 2005)
NewcombÕs Problem (Stanovich & West, 1999; Toplak & Stanovich, 2002)	


Is Mindware Available to Carry Out Override? (Subject Parameter #1)
Yes
No
Heuristic Response Path #1
Does Subject Detect the Need to Override the Heuristic Response? (Subject Parameter #2)
Yes
No
Heuristic Response Path #2
Is Sustained Inhibition or Sustained Decoupling Necessary to Carry Out Override? (Task Factor)
Yes
No
System 2 Response
Does Subject Have Decoupling Capacity to Sustain Override? (Subject Parameter #3)
Yes
No
Heuristic Response Path #3
System 2 Response

Tasks in these Experiments and in Other Studies that do and do not Show Associations with Cognitive Ability

Tasks/Effects that Fail to Correlate with Cognitive Ability	Tasks/Effects that Correlate with Cognitive Ability
Noncausal Baserate Usage (Experiment 1; Stanovich & West, 1998c, 1999)	Causal Baserate Usage (Stanovich & West, 1998c, 1998d)
Conjunction Fallacy Between-Subjects (Experiment 1)	Outcome Bias (Experiment 1; Stanovich & West, 1998c)
Framing Between-Subjects (Experiment 1)	Framing Within-Subjects (Frederick, 2005; Parker & Fischhoff, 2005; Stanovich & West, 1998b, 1999)
Anchoring Effect (Experiment 1)	Denominator Neglect (Experiment 8; Kokis et al., 2002)
Evaluability Less is More Effect (Experiment 2)	Probability Matching (Experiment 8; West & Stanovich, 2003)
Proportion Dominance Effect (Experiment 2)	Hindsight Bias (Stanovich & West, 1998c)
Sunk Cost Effect (Experiment 3; Parker & Fischhoff, 2005)	Ignoring P(D/NH) (Stanovich & West, 1998d, 1999)
Risk/Benefit Confounding (Experiment 4)	Covariation Detection (Stanovich & West, 1998c, 1998d; Sa et al., 1999)
Omission Bias (Experiment 5)	Belief Bias in Syllogistic Reasoning (Experiment 8; Stanovich & West, 1998c)
Perspective Bias (Experiment 5)	Belief Bias in Modus Ponens (Experiment 8)
Certainty Effect (Experiment 6)	Informal Argument Evaluation (Experiment 8; Stanovich & West, 1997)
WTP/WTA Difference (Experiment 6)	Four Card Selection Task (Experiment 8; Stanovich & West, 1998a)
Myside Bias - Between and Within-S (Experiment 7; Stanovich & West, 2007)	EV Maximization in Gambles (Frederick, 2005; Benjamin & Shapiro, 2005)
NewcombÕs Problem (Stanovich & West, 1999; Toplak & Stanovich, 2002)	