PA202: Cognitive Biases when Evaluating Academic Progress Monitoring Data

Adelle K. Sturgell & Dr. Ethan R. Van Norman

LEHIGH UNIVERSITY

Abstract

School psychologists' instructional decisions can be impacted by cognitive biases even within response-to-intervention models. While decision-making in conjunction with psychoeducational reports has been well-researched, the impacts of cognitive biases on progress monitoring decisions are less understood. Results of this study will highlight inhibitors to accurate data-based decision making. Session attendees will receive information on cognitive bias identification in their own practices, ensuring that their students will in fact continue to respond to intervention.

LEHIGH UNIVERSITY

Aligns with NASP Practice Model Domain 1

LEHIGH UNIVERSITY

Introduction

- Research in Decision Making
  - School psychologists make a wide variety of decisions of varying levels of importance
  - The accuracy of some of these decisions can be poor (Maki et al., 2021)

Screening
- Which individual students are likely to need help?
- How well is our system doing at meeting the needs of our students?

Instructional Planning
- How should this student be instructed?
- What supports does this student need to be successful?

Resource Allocation
- How should we distribute our time and effort?

Special Education Eligibility
- Does this student qualify for special education services?

Progress Monitoring
- Is this student improving at an acceptable rate?
- Do we need to make adjustments to their instructional program?

Program Evaluation
- Are our instructional practices effective?

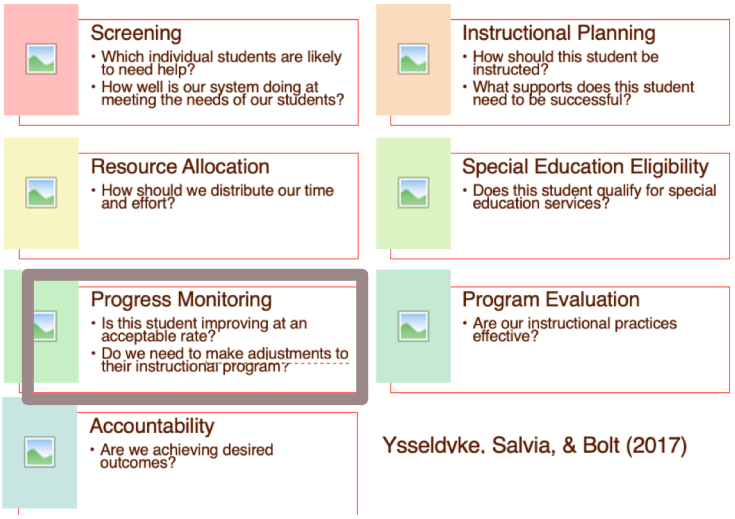
Accountability
- Are we achieving desired outcomes?

Ysseldyke, Salvia, & Bolt (2017)

LEHIGH UNIVERSITY

Introduction

- Research in Decision Making
  - Shift from the sole usage of cognitive aptitude tests (Maki et al., 2015) to response to scientifically-based interventions/measures that can be repeatedly administered over time (Benson et al., 2020)
  - Measures match students to interventions based upon specific needs (Fletcher & Vaughn, 2009), but the methods are not perfect (Reynolds & Shaywitz, 2009)

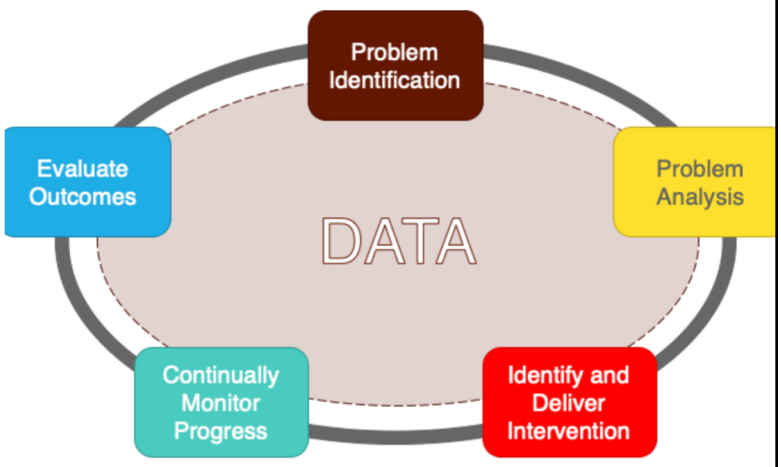


Ysseldvke, Salvia, & Bolt (2017)

LEHIGH UNIVERSITY

Introduction

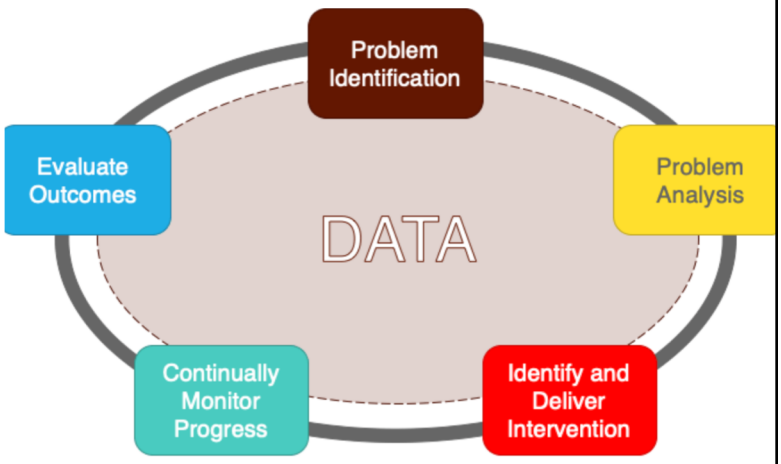
- Decision Making, Problem Solving & MTSS
  - When making decisions, follow the process →
  - Many of the decisions we make come before the decision on special education eligibility



LEHIGH UNIVERSITY

Introduction

- Decision Making, Problem Solving & MTSS
  - Is the student improving appropriately relative to the level of support they are receiving?
  - At what point does the student need more intensive instruction?
  - Amount of data collected influences the reliability of student growth estimates (Christ, 2006)
  - Proactive decision-making is key (King & Coughlin, 2016)



LEHIGH UNIVERSITY

Introduction

- Cognitive Biases
  - Unconscious errors that can occur when we make challenging decisions (Davidow & Levinson, 1993)
  - Can be attributed to various reasoning mistakes
  - Examples:
    - Dunning-Kruger Effect: people less competent at tasks overestimate their ability to perform that task, while more competent people tend to underestimate their capabilities (Kruger & Dunning, 1999)
    - Law of the Instrument: people choose the option they are most comfortable with even if the option is not appropriate for the task (Lee, 2016)

LEHIGH UNIVERSITY

Introduction

- Cognitive Biases
  - Unconscious errors that can occur when we make challenging decisions (Davidow & Levinson, 1993)
  - Can be attributed to various reasoning mistakes
  - Examples:
    - Dunning-Kruger Effect
    - Law of the Instrument
  - Sunk-Cost Fallacy - our focus for today
    - Continuing a course of action because of the time we have already invested, regardless of whether or not we should actually continue
    - Can happen to anyone:
      - Business students (Staw, 1976)
      - School psychologists! (Wilcox & Schroeder, 2015)
    - Difference between making the decisions yourself vs. taking over when someone else has already made the decisions

LEHIGH UNIVERSITY

Purpose & Research Questions

Purpose: to explore whether the sunk-cost fallacy influences data-based decision making processes when evaluating CBM or oral reading progress monitoring data

Research Questions:

(1) How many data points do visual analysts tend to collect before they make a treatment decision?

(2) To what degree is the number of data points required to make a decision influenced by whether participants choose the intervention students received?

(3) To what degree is the accuracy of treatment decisions influenced by whether participants choose the intervention students receive?

LEHIGH UNIVERSITY

Method

- Participants
  - 30 participants sampled from professional school psychology organizations
  - Sample demographics reflect many characteristics seen in the field

Demographic Information of Study Participants

	No Choice		Choice		Entire Sample	
Characteristic	*N*	%	*N*	%	*N*	%
Sex						
Male	3	20	5	33	8	27
Female	12	80	10	67	22	73
Race/Ethnicity						
White	12	80	15	100	27	90
Hispanic	1	7	0	0	1	3
Black	1	7	0	0	1	3
More than one race	1	7	0	0	1	3
Degree Level						
Bachelor's	0	0	2	13	2	7
Master's	5	33	2	13	7	23
Specialist	7	47	7	47	14	47
Doctorate	3	20	4	27	7	23
Geographic Location						
Northeast	8	53	7	47	15	50
Southeast	1	7	3	20	4	13
Midwest	3	20	2	13	5	17
Southwest	2	13	0	0	2	7
West	1	7	3	20	4	13
Occupation						
School Psychologist	13	87	10	67	23	77
Teacher	0	0	2	13	2	7
Administrator	0	0	1	7	1	3
Faculty	2	13	2	13	4	13

LEHIGH UNIVERSITY

Method – Materials

- Participants viewed 36 CBM-R Progress Monitoring Graphs
  - Schedule – One observation per week
- Number of Initial Observations Plotted
  - 4, 6, or 8
- Patterns of ”True-Trend”
  - Initial Observations vs. Remaining Observations
  - Low-Low; Low-High; High-High; High-Low

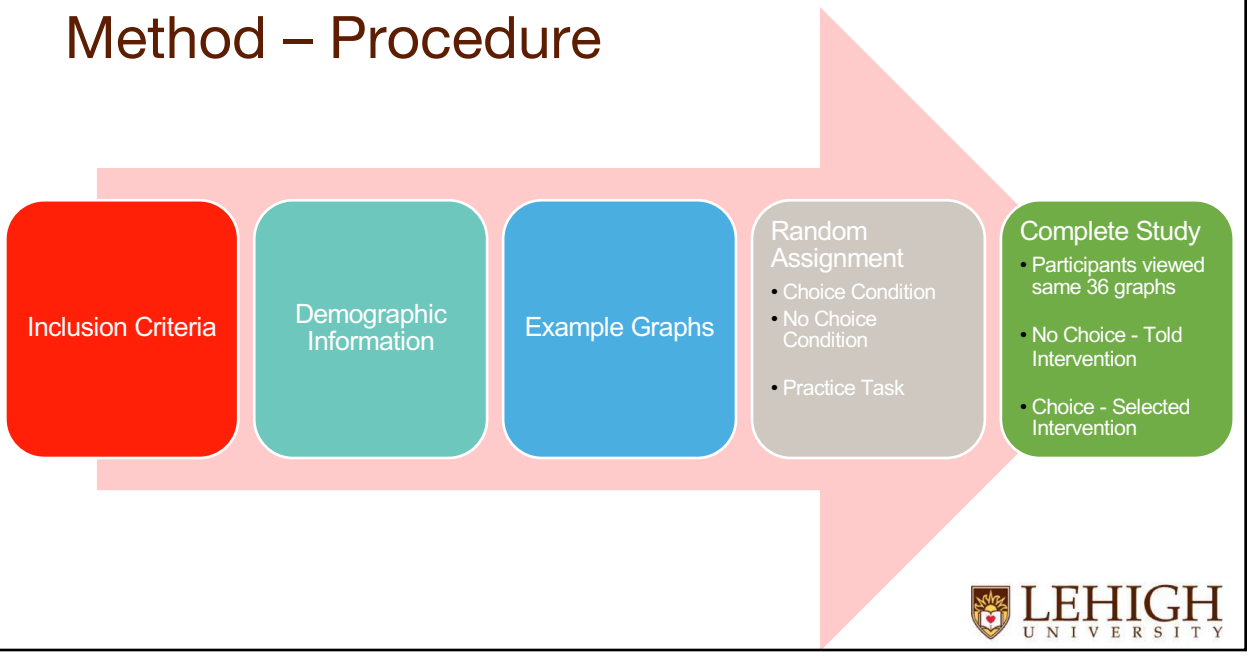
LEHIGH UNIVERSITY

Method – Materials

- Participants viewed 36 CBM-R Progress Monitoring Graphs
  - Schedule – One observation per week
- Number of Initial Observations Plotted
  - 4, 6, or 8
- Patterns of "True-Trend"
  - Initial Observations vs. Remaining Observations
  - Low-Low; Low-High; High-High; High-Low
- 12 Unique Conditions → 3 Graphs per Condition → 36 Graphs

LEHIGH UNIVERSITY

Method – Procedure



LEHIGH UNIVERSITY

Method – Analysis

- Dependent Variable 1 – Number of Data Points Requested
  - Descriptive Analyses
  - Inferential (Linear Mixed Effects Regression)

- Dependent Variable 2 – Accuracy of Recommendation
  - Low-Low & High-Low – Make a Change
  - High-High & Low-High – Maintain
  - 0 – incorrect response 1-correct response

  - Descriptive
  - Inferential (Generalized Linear Mixed Effects Regression)

LEHIGH UNIVERSITY

Results

Summary of Participant Self-Reported Background with Curriculum Based Measurement and Academic Interventions

	No Choice		Choice		Whole Sample	
	M	*SD*	*M*	*SD*	*M*	*SD*
Age	41.6	11.27	35.6	7.41	38.6	9.86
Years of Experience	13.50	10.89	8.20	4.99	10.85	8.76
Years Using CBM	12.30	8.08	9.20	6.13	10.6	7.27
*Confidence in Accurately Collecting CBM Data	87.70	18.99	82.90	24.00	85.30	17.97
*Confidence in Using CBM Data to make Instructional Decisions	87.20	10.25	83.70	18.78	85.47	14.97
*Emphasis Your Training Program Placed on Using CBM	79.80	23.55	77.30	25.52	78.57	24.16
*Familiarity with Interventions for ORF	79.80	23.90	79.30	23.77	79.53	23.43
*Emphasis your training program placed on implementing academic interventions	71.40	29.45	79.10	20.76	75.23	25.34

* Participants were asked to give their responses on a scale from 1 (low / not confident) to 100 (high / confident).
Note. CBM – curriculum based measurement; ORF – oral reading fluency

LEHIGH UNIVERSITY

Results – Number of Data Points

- Overall, participants requested an additional *M* = 2.59 (*SD* = 3.22) observations before making a decision

- Initial Observations
  - 4: *M* = 4.12 (*SD* = 3.43)
  - 6: *M* = 2.38 (*SD* = 3.28)
  - 8: *M* = 1.27 (*SD* = 2.11)

- True Trend
  - Low-Low: 2.62 (*SD* = 3.60)
  - High-High: 1.91 (*SD* = 2.71)
  - Low – High 2.53 (*SD* = 2.25)
  - High – Low: 3.29 (*SD* = 3.89)

LEHIGH UNIVERSITY

No Choice:
M = 2.27 (*SD* = 3.03)

Choice:
M = 2.90 (*SD* = 3.37)

Descriptive Results for the Number of Observations Requested Before Making a Treatment Decision

		No Choice		Choice	
Trend Condition	Initial Observations	*M*	*SD*	*M*	*SD*
High High	4	4.02	3.13	3.58	2.53
	6	1.36	2.02	1.58	3.27
	8	0.42	0.84	0.53	1.14
High Low	4	4.96	4.40	4.13	4.25
	6	3.87	3.78	2.80	3.42
	8	2.49	3.73	1.49	2.69
Low High	4	4.16	2.67	3.62	2.71
	6	1.80	1.52	1.13	1.44
	8	2.38	1.45	2.07	1.83
Low Low	4	4.62	3.29	3.84	3.90
	6	4.36	4.84	2.13	3.03
	8	0.42	0.99	0.36	0.71

Note. High referred to a true trend equal to 3.00 words read correct per minute improvement per week. Low referred to a true trend equal to 0.00 words read correct per minute improvement per week. The first descriptor referred to the true trend for initial observations shown to visual analysts and the second descriptor referred to true trend for the remaining observations requested by the participant in the series.

Linear Mixed Effects Regression Results for Number of Data Points Requested Across Study Conditions

	Baseline A		Model 1		Model 2		Model 3	
Fixed Effects	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	3.44	0.34	4.12	0.32	3.44	0.34	3.14	0.45
6 Observations			-1.74	0.19	-1.74	0.19	-1.74	0.19
8 Observations			-2.85	0.19	-2.85		-2.85	0.19
High Low Trend					1.37	0.22	1.37	0.22
Low High Trend					0.61	0.22	0.61	0.22
Low Low Trend					0.71	0.22	0.71	0.22
No Choice Group							0.63	0.58
Random Effects	*SD*		*SD*		*SD*		*SD*	
Intercept	1.56		1.57		1.57		1.54	
Residual	2.81		2.55		2.50		2.50	
Model Fit	Deviance	AIC	Deviance	AIC	Deviance	AIC	Deviance	AIC
	5374.27	5380.27	5168.00	5178.00	5127.85	5143.85	5126.69	5144.69
	χ^2(df)	*p*	χ^2(df)	*p*	χ^2(df)	*p*	χ^2(df)	*p*
	-	-	206.27 (2)	<.001	40.16 (3)	<.001	1.16 (1)	.283

Note. Fixed effect predictors were added as a categorical covariates. The referent (intercept) conditions include: 4 initial observations, High High Trend, and visual analysts in the Choice condition. AIC – Akaike Information Criterion.

Results – Decision Accuracy

- Across all conditions, 82.96% of recommendations were correct.
- Number of initial observations *did not* influence accuracy $\chi^2(2) = 1.94$, $p = .379$
  - 4: 83%
  - 6: 85%
  - 8: 81%
- True trend condition *did* influence accuracy $\chi^2(3) = 90.77$, $p < .001$
  - Low Low: 85%
  - High High: 95%
  - High Low: 65%
  - Low High: 87%

Results – Decision Accuracy

- Group assignment was related to accuracy $\chi^2(1) = 15.73$, $p < .001$

Choice 78% vs.

No Choice 88%

Summary of the Percentage of Incorrect and Correct Recommendations for Visual Analysts in No Choice and Choice Groups Teased Apart by Trend Condition

	No Choice		Choice	
Trend Condition	Incorrect	Correct	Incorrect	Correct
High High	0.00	100.00	10.37	89.63
High Low	28.89	71.11	40.74	59.26
Low High	10.37	89.63	16.30	83.70
Low Low	10.37	89.63	19.26	80.74

LEHIGH UNIVERSITY

Generalized Linear Mixed Effects Regression Results for Log-Odds of a Correct Decision for Visual Analysts

	Baseline B		Model 4		Model 5		Model 6	
Fixed Effects	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	1.81	0.19	1.80	0.22	3.34	0.34	2.93	0.38
6 Observations			0.18	0.21				
8 Observations			-0.13	0.20				
High Low Trend					-2.59	0.32	-2.59	0.32
Low High Trend					-1.13	0.34	-1.13	0.34
Low Low Trend					-1.27	0.33	-1.27	0.33
No Choice Group							0.83	0.39
Random Effects	*SD*		*SD*		*SD*		*SD*	
Intercept	0.87		0.87		0.99		0.91	
Model Fit	Deviance	AIC	Deviance	AIC	Deviance	AIC	Deviance	AIC
	927.14	931.14	924.96	932.97	827.25	837.25	758.93	
	χ^2(df)	*p*	χ^2(df)	*p*	χ^2(df)*	*p*	χ^2(df)	*p*
	-	-	2.17 (2)	.338	99.89 (3)	<.001	4.40 (1)	.036

*In comparison to Baseline B

Note. Fixed effect predictors were added as a categorical covariates. The referent (intercept) conditions include: 4 initial observations, High High Trend, and visual analysts in the Choice condition. AIC – Akaike Information Criterion.

Results – Decision Accuracy

$$P = \frac{e^{log-odds}}{1 + e^{log-odds}}$$

True Trend Condition	No Choice	Choice
Low – Low	.92	.84
High - High	.98	.95
High – Low	.76	.58
Low - High	.93	.86

LEHIGH UNIVERSITY

Discussion

Purpose: to explore whether the sunk-cost fallacy influences data-based decision making processes when evaluating CBM or oral reading progress monitoring data

Research Questions:

(1) How many data points do visual analysts tend to collect before they make a treatment decision?

(2) To what degree is the number of data points required to make a decision influenced by whether participants choose the intervention students received?

(3) To what degree is the accuracy of treatment decisions influenced by whether participants choose the intervention students receive?

LEHIGH UNIVERSITY

Discussion

Research Questions:

(1) How many data points do visual analysts tend to collect before they make a treatment decision?

- Visual Analysts tended to make a decision at around 8-10 observations
- When data showed consistent patterns fewer observations were needed
- Steep initial performance followed by deceleration required the most observations

LEHIGH UNIVERSITY

Discussion

Research Questions:

(2) To what degree is the number of data points required to make a decision influenced by whether participants choose the intervention students received?

- Statistically controlling for other characteristics having a choice vs. not having a choice did **not** influence the number of additional observations requested.

LEHIGH UNIVERSITY

Discussion

Research Questions:

(3) To what degree is the accuracy of treatment decisions influenced by whether participants choose the intervention students receive?

- Number of initial observations did not influence accuracy of decisions
  - Aligns with findings from RQ 1 (Newell & Christ, 2017)
- Consistency of true trend influenced the accuracy of decisions
- Visual analysts that did not choose the intervention were more accurate statistically controlling for other characteristics

LEHIGH UNIVERSITY

Discussion – Implications

- Explicit practice visually analyzing graphs (Wolf & Slocum, 2015) specifically CBM graphs (Espin et al., 2017)
  - Need to identify cognitive biases to inform these efforts
- Distinction between standard-treatment protocol and problem-solving approaches to RTI
  - May have a propensity to over-identify interventions we select as effective
- Rapid growth followed by deceleration is **not** unheard of (Van Norman & Parker, 2018)
  - Some students struggle to maintain acquired skills after showing initial improvement (Nelson et al., 2018)

LEHIGH UNIVERSITY

Discussion - Limitations

- Sample Size
  - 30 participants
  - Participants were people who were fairly familiar with the assigned task
  - Limited sample size can complicate generalizability
- Use of a Vignette
  - Vignette usage allows for higher experimental control - we could isolate the impacts of progress monitoring on decision making
  - If the study were to be conducted in a more naturalistic setting, participants would have been able to utilize background and other forms of information that impact decision making
- Other Contexts and Considerations

LEHIGH UNIVERSITY

Discussion – Future Directions

- Potential for a team-based decision making study
  - In this study, participants worked individually, but these types of decisions are often made within team processes
  - Educators could initially evaluate on their own and then compare answers with other participants
- Manipulation of other graph features
  - Only limited features were manipulated in this study in order to make the tasks more manageable
  - Manipulate graph features (i.e., scaling of x- or y-axes)
  - Manipulate features of the plotted data (i.e., level of measurement error)

LEHIGH UNIVERSITY