VANDERBILT
PEABODY COLLEGE

PEABODY research institute

Academically Relevant Measures of Self-Regulation:

Development and Validation of Assessments for Preschool Children

Kimberly A. Turner, Mark W. Lipsey, Dale C. Farran, Nianbo Dong,
Sandra J. Wilson, Mary Wagner Fuhs, Elizabeth A. Vorhaus, & Deanna N. Meador
Peabody Research Institute, Vanderbilt University

Society for Research on Educational Effectiveness Spring 2012 Conference
March 8, 2012

Project funded by the Institute of Education Sciences (R305A090079). First author is supported by an Institute of Education Sciences Postdoctoral Fellowship (R305B100016)

Research Team

Principal Investigator

- Mark W. Lipsey

Co-Principal Investigators

- Dale C. Farran
- Sandra J. Wilson

Research Coordinators

- Elizabeth A. Vorhaus
- Deanna N. Meador

Research Analysts

- Nianbo Dong
- Mary Wagner Fuhs
- Kimberly A. Turner

Research Assistants

- Ashley G. Keene
- Jennifer L. Norvell

Predoctoral Fellows

- Karen Anthony
- Lydia Bentley
- Amy Holmes
- Katherine Newman
- Cathy Yun

Multiple child assessors and PRI support staff

Project funded by the Institute of Education Sciences (R305A090079)

First author support by IES Postdoctoral Fellow ship (R305B100016)

School Readiness and Self-Regulation (SR)

- Growing evidence that young children's self-regulation is positively related to school success
  - Robust and consistent relations found when using with teacher reports of children's self-regulation
- Direct assessment of young children's self-regulation are necessary for evaluation of intervention and curriculum
  - Need measures that align with teacher ratings
  - Need measures that predict subsequent achievement
  - Need measures whose gains relate to gains in achievement

Purpose of Study

- Construct a direct assessment battery of SR for preschool children that:
  - Can be easily administered in preschool settings
  - Shows joint variation with other SR measures
  - Exhibits variability and gains from the beginning to the end of preschool
  - Predicts academic achievement, including gains in achievement (predictive validity)
  - Converges with teacher ratings of classroom self-regulation (convergent validity)

Significance of Study

- Validated battery of SR measures is critical for educational research
  - Enhance the quality of descriptive research
  - Allow for screening of young children's SR as indicator of school readiness
  - Provide effective methods for assessing improvements in SR that result from classroom practices and interventions

Operational Framework

- **Sustained Focus:** Conscious detection and continued attention (Posner & Rothbart, 2000)
- **Effortful Control:** Suppression of emotional and behavioral reactivity (Rothbart & Ahadi, 1994)
- **Inhibitory Control:** Volitional inhibition of a prepotent cognitive response (Diamond, 1990)
- **Attention Shifting:** Adaptation between distinct but related mental sets (Zelazo, Frye, & Rapus, 1996)
- **Working Memory:** Active maintenance and manipulation of information (Baddeley & Hitch, 1974)

Assessment Selection

- **Sustained Focus**
  - **Copy Design** (Davie, Butler, & Goldstein, 1972; Osborn, Butler, & Morris, 1984)
  - **Kansas Reflection-Impulsivity Scale for Preschoolers (KRISP)** (Wright 1971)
- Effortful Control
- Inhibitory Control
- Attention Shifting
- Working Memory

Assessment Selection

- Sustained Focus
- **Effortful Control**
  - **Whisper** (Kochanska, Murray, Jacques, Koenig, & Vandegeest, 1996)
  - **Turtle-Rabbit** (Kochanska, Murray, Jacques, Koenig, & Vandegeest, 1996)
- Inhibitory Control
- Attention Shifting
- Working Memory

Assessment Selection

- Sustained Focus
- Effortful Control
- **Inhibitory Control**
  - **Head Toes Knees Shoulders (HTKS)** (McClelland et al., 2007)
  - **Peg Tapping** (Diamond & Taylor, 1996)
  - **Spatial Conflict Arrows** (Blair & Willoughby, 2006; Willoughby et al., in press)
- Attention Shifting
- Working Memory

Assessment Selection

- Sustained Focus
- Effortful Control
- Inhibitory Control
- **Attention Shifting**
  - **Dimensional Change Card Sort (DCCS)** (Zelazo, 2006)
- Working Memory

Assessment Selection

- Sustained Focus
- Effortful Control
- Inhibitory Control
- Attention Shifting
- **Working Memory**
  - **Operation Span** (Blair & Willoughby, 2006)
  - **Backward Digit Span** (Davis & Pratt, 1996)

Setting and Participants

- 535 PreK children from 58 classrooms within 38 schools/centers across 4 school systems
  - Schools
    - United Way and Public Pre-K Schools
    - Urban and Suburban School Systems
  - Children
    - Mean age 4 years, 6 months at beginning of PreK
    - 52% male
- 488 children located and completed end of Kindergarten assessments (91% retention)

Criterion Measures

- Predictive Validity: Academic Achievement
  - Woodcock-Johnson III: Applied Problems, Quantitative Concepts, Picture Vocabulary, Letter-Word Identification, and Oral Comprehension (Woodcock, McGrew, Mather, 2001)
- Convergent Validity: Child Behavioral Ratings
  - Cooper- Farran Behavioral Ratings Scales (CFBRS): Work-Related Skills subscale (Cooper & Farran, 1988)
  - Temperament Assessment Battery for Children (TABC): Distractibility and Persistence subscales (Martin, 1988)
  - Children's Behavioral Questionnaire (CBQ): Impulsivity and Attention Shifting subscales (Rothbart, Ahadi, Hershey, & Fisher, 2001)

Procedures

- Three time points:
  - Beginning of PreK (T1)
  - End of PreK (T2)
  - End of Kindergarten (T3)
- Direct Assessment
  - SR tasks
  - Academic Achievement Measures
- Child Behavioral Ratings
  - Completed by teachers
  - Paper and Electronic Administration

Results: Joint Variation

Loadings from 1-factor PFA Solution PFA

Direct Assessment	Time 1	Time 2
Peg Tapping	0.70	0.70
HTKS	0.64	0.68
DCCS	0.53	0.54
Backward Digit Span	0.40	0.46
Copy Design	0.51	0.53
KRISP Accuracy	0.65	0.61
KRISP Reaction Time	0.39	0.35
Turtle-Rabbit Accuracy	0.35	0.33
Turtle-Rabbit Reaction Time	0.41	0.46
Whisper Task	0.41	0.32
Spatial Conflict	0.50	0.39
Operation Span	0.40	0.34

Results: Developmental Appropriateness

- Differences in T1 & T2 performance (Effect Size)

Direct Assessment	*d*
Peg Tapping	.62*
HTKS	.55*
DCCS	.52*
Backward Digit Span	.44*
Copy Design	.55*
KRISP Accuracy	.67*

Direct Assessment	*d*
KRISP Reaction Time	.37*
Turtle-Rabbit Accuracy	.00
Turtle-Rabbit Reaction Time	.41*
Whisper Task	.37*
Spatial Conflict	.36*
Operation Span	.31*

*Significant difference between T1 and T2 performance, *t*s ≥ 5.89, *p*s < .01

Results: Developmental Appropriateness

- Correlations between T1 & T2 performance

Direct Assessment	*r*
Peg Tapping	.62*
HTKS	.66*
DCCS	.38*
Backward Digit Span	.46*
Copy Design	.60*
KRISP Accuracy	.56*

Direct Assessment	*r*
KRISP Reaction Time	.12*
Turtle-Rabbit Accuracy	.20*
Turtle-Rabbit Reaction Time	.44*
Whisper Task	.35*
Spatial Conflict	.33*
Operation Span	.39*

*Significant difference between T1 and T2 performance, *ps* < .01

Results: Predictive Validity

Multilevel Models with T1 SR measures entered simultaneously

Direct Assessment	SR T1 → AA T2	SR T1 → AA T3
Peg Tapping	.14**	.14**
HTKS	.18**	.15**
DCCS	.14**	.16**
Backwards Digit Span	.19**	.16**
Copy Design	.11**	.08*
KRISP Accuracy	.15**	.23**
KRISP Reaction Time	-.01	-.01
Turtle-Rabbit Accuracy	.01	.03
Turtle-Rabbit Reaction Time	.08*	.04
Whisper Task	.14**	.11**
Spatial Conflict	.05	.03
Operation Span	.02	.03

Note. Estimates are standardized values and can be interpreted as partial correlation coefficients.
**$p < .01$, *$p < .05$

Results: Predictive Validity

Direct Assessment	SR Gain →AA T1 T2 Gain
Prior Achievement	.75**
Peg Tapping	.07**
HTKS	.03
DCCS	.06**
Backwards Digit Span	.10**
Copy Design	.04†
KRISP Accuracy	.04†
KRISP Reaction Time	.01
Turtle-Rabbit Accuracy	.05†
Turtle-Rabbit Reaction Time	.06**
Whisper Task	.06**
Spatial Conflict	.03
Operation Span	.03

Multilevel Models with Residualized SR Gains predicting T2 achievement controlling for prior achievement.

Estimates are standardized values and can be interpreted as partial correlation coefficients.
$^{**}p < .01$, $^{*}p < .05$; $^{\dagger}p < .10$

Results: Predictive Validity

Direct Assessment	SR Gain →T1 T3 AA Gain	SR Gain →T2 T3 AA Gain
Prior Achievement	.69**	.78**
Peg Tapping	.04	-.01
HTKS	.12**	.11**
DCCS	.03	.00
Backwards Digit Span	.11**	.05†
Copy Design	.04	.02
KRISP Accuracy	.04	.01
KRISP Reaction Time	.01	.00
Turtle-Rabbit Accuracy	.00	-.03
Turtle-Rabbit Reaction Time	.09**	.07**
Whisper Task	-.02	-.06*
Spatial Conflict	.02	.00
Operation Span	.00	-.02

Note. Estimates are standardized values and can be interpreted as partial correlation coefficients.
**$p < .01$, *$p < .05$; †$p < .10$

Results: Convergent Validity

Average T1 & T2 SR Correlations w/ Child Behavioral Ratings

	CFBRS	TABC		CBQ	
Direct Assessment	WRS	Distract	Persist	Impulse	Atten Shift
Peg Tapping	.41**	.39**	.36**	.31**	.18**
HTKS	.38**	.36**	.34**	.29**	.14**
DCCS	.26**	.26**	.26**	.19**	.13**
Backwards Digit Span	.22**	.19**	.19**	.14**	.04
Copy Design	.33**	.30**	.31**	.19**	.10*
KRISP Accuracy	.38**	.35**	.35**	.28**	.18**
KRISP Reaction Time	.21**	.16**	.18**	.15**	.02
Turtle-Rabbit Accuracy	.24**	.27**	.22**	.23**	.15**
Turtle-Rabbit Reaction Time	.25**	.22**	.17**	.16**	.05
Whisper Task	.22**	.23**	.20**	.16**	.09*
Spatial Conflict	.22**	.24**	.24**	.18**	.18**
Operation Span	.21**	.18**	.15**	.13**	.07

**p < .01, *p <.05

Refinement Selection

- Copy Design, DCCS, HTKS, KRISP Accuracy, Peg Tapping
  - Joint Covariance
  - Developmentally Appropriate
  - Entering Performance Predictive of Future Achievement
  - Gains on Performance Predictive of Achievement Gains
  - Convergent Validity with Teacher Ratings
- Backward Digit Span
  - Developmentally Appropriate
  - Entering Performance Predictive of Future Achievement
  - Gains on Performance Predictive of Achievement Gains

Predictive Validity of Refined Battery

- Data Reduction via PCA of Six Selected Assessment (1-Factor Solution)

	Time 1	Time 2
Peg Tapping	.77	.76
HTKS	.75	.76
DCCS	.64	.66
Backward Digit Span	.53	.58
Copy Design	.62	.63
KRISP Accuracy	.67	.65

Predictive Validity of Refined Battery

- Multilevel Models Predicting Academic Achievement

	SR	Prior Achievement
SR T1 → AA T2	.71**	--
SR T1 → AA T3	.68**	--
SR Gain → AA T1 T2 Gain	.15**	.84**
SR Gain → AA T1 T3 Gain	.13**	.77**
SR Gain → AA T2 T3 Gain	.02	.81**

Note. Estimates are standardized values and can be interpreted as partial correlation coefficients.
**$p < .01$

Conclusion

- A relatively clear identification of the SR assessments that best meet our criteria
- SR performance at the beginning of PreK predicted achievement at the end of both PreK and Kindergarten
- Gains on the SR assessments across the PreK year predicted PreK and Kindergarten gains in achievement
- Performance on SR assessment converge with Teacher's ratings of SR witnessed in the classroom

Future Directions

- Develop scoring system to combine individual scores into single composite score
- Cross-Validation of refined battery, including establish test-retest reliability
- Create and validate parallel child behavioral rating scale

For more information please contact

Kimberly A. Turner
IES Postdoctoral Research Fellow

E-mail: kimberly.turner@vanderbilt.edu
http://peabody.vanderbilt.edu/pri.xml

Peabody Research Institute
Vanderbilt University
Box 0181 GPC, 230 Appleton Place
Nashville, TN 37203-5721
(615) 322-8015
(615) 322-0293 (F)