

**Table S1**  
**Component Ratings for Strength of Study Design**

Author	Design	IV	DV	IOA	IOA more than 80%	Fidelity	Three attempts	Points per phase	Overall Design Rating
Benish & Bramlett (2011)	MBD	Yes	Yes	No	No	Yes	Yes	3	Does not meet
Burke et al. (2004)	ABAB	Yes	Yes	Yes	Yes	Yes	No	1	Does not meet
	MBD	Yes	Yes	Yes	Yes	Yes	Yes	4	Meets w/reservations
Chan & O'Reilly (2008)	Multiprobe	Yes	Yes	Yes	Yes	Yes	No	3	Does not meet
Crozier & Tincani (2007)	ABAB	Yes	Yes	Yes	Yes	Yes	Yes	4	Meets w/reservations
	ABCACBC	Yes	Yes	Yes	Yes	Yes	Yes	2	Does not meet
	ABAB	Yes	Yes	Yes	Yes	Yes	Yes	5	Meets standards
Hsu et al. (2012)	MBD	Yes	Yes	Yes	Yes	Yes	Yes	3	Meets w/reservations
Ivey et al. (2004)	ABAB	Yes	Yes	Yes	Yes	Yes	Yes	4	Meets w/reservations
	ABAB	Yes	Yes	Yes	Yes	Yes	Yes	4	Meets w/reservations
Kuoch & Mirenda (2003)	ABA	Yes	Yes	Yes	Yes	Yes	No	5	Does not meet
	ABA	Yes	Yes	Yes	Yes	Yes	No	8	Does not meet
Leaf et al. (2012)	MBD	Yes	Yes	Yes	Yes	Yes	Yes	3	Meet w/reservations
Lorimer et al. (2002)	ABAB	Yes	Yes	Yes	Yes	No	Yes	3	Meet w/reservations
Schneider & Goldstein (2010)	MBD	Yes	Yes	Yes	Yes	Yes	Yes	9	Meets standards
Soenksen & Alper (2006)	MBD	Yes	Yes	Yes	Yes	No	Yes	3	Meets w/reservations
Wright & McCathren (2012)	MBD	Yes	Yes	Yes	Yes	Yes	Yes	5	Meets standards

**Table S2**  
**Component Ratings for Evidence of Experimental Control**

Author	Design	Stable baseline	Overlapping data	Immediacy	Consistency	Functional Rel.	Evidence Rating
Benish & Bramlett (2011)	MBD	No	Yes	Yes	No	No	No evidence
	MBD	No	Yes	Yes	No	Yes	Moderate
	MBD	No	Yes	Yes	No	Yes	Moderate
Burke et al. (2004)	ABAB	No	Yes	No	No	No	No evidence
	MBD	No	Yes	Yes	Yes	Yes	Moderate
Chan & O'Reilly (2008)	Multi-probe	No	Yes	Yes	Yes	No	No evidence
Crozier & Tincani (2007)	ABAB	No	Yes	Yes	Yes	Yes	Moderate
	ABCACBC	Yes	Yes	Yes	Yes	No	No evidence
	ABAB	Yes	Yes	Yes	Yes	Yes	Moderate
	ABAB	No	Yes	No	No	No	No evidence
Hsu et al. (2012)	MBD	No	Yes	No	No	No	No evidence
Ivey et al. (2004)	ABAB	No	Yes	No	No	No	No evidence
	ABAB	No	Yes	No	No	No	No evidence
Kuoch & Mirenda (2003)	ABA	No	Yes	Yes	No	No	No evidence
	ABA	No	Yes	Yes	No	No	No evidence

Leaf et al. (2012)	MBD – Fig. 2-4	Yes	Yes	No	No	No	No evidence
Lorimer et al. (2002)	ABAB	No	Yes	Yes	Yes	Yes	Moderate
	ABAB	No	Yes	Yes	Yes	Yes	Moderate
Schneider & Goldstein (2010)	MBD	No	Yes	Yes	Yes	Yes	Moderate
Soenksen & Alper (2006)	MBD	Yes	Yes	Yes	Yes	Yes	Moderate
	MBD	Yes	Yes	Yes	Yes	Yes	Moderate
Wright & McCathren (2012)	MBD	Yes	Yes	No	No	No	No evidence
	MBD	No	Yes	No	No	No	No evidence

**Table S3**  
**Meta-regression of Log Response Ratio-increasing effect size estimates**

Predictor	Coefficient	
	Estimate (SE)	95% CI
Challenging behavior (average LRRi)	1.02 (0.22)	[-0.11, 2.15]
Pro-social behavior (average LRRi)	0.83 (0.51)	[-1.11, 2.77]
<i>WWC design rating<sup>a</sup></i>		

Meets standards with reservations	0.01 (0.51)	[-1.37, 1.38]
Does not meet standards	0.35 (0.55)	[-1.14, 1.85]
<b><i>Participant age in years<sup>b</sup></i></b>		
Six or older	-0.11 (0.56)	[-2.05, 1.84]
<b><i>Participant diagnosis<sup>c</sup></i></b>		
No diagnosed disability	0.18 (0.39)	[-0.91, 1.26]
<b><i>Interventionist and primary data collector<sup>d</sup></i></b>		
Same	0.14 (0.32)	[-0.66, 0.94]

---

Notes:

<sup>a</sup> Reference category is meets standards without reservations.

<sup>b</sup> Reference category is five or younger.

<sup>c</sup> Reference category is diagnosed disability.

<sup>d</sup> Reference category is Interventionist was not the primary data collector.

**Additional discussion of Table S3**

Table S3 reports the results of a joint meta-regression, controlling for additive effects of each of the four potential moderators while also differentiating between challenging behavior and prosocial behavior outcomes. After controlling for all four moderators, the average LRRi effect sizes for challenging behavior and prosocial behavior were reduced and were imprecisely estimated, so that they were not statistically distinguishable from null. The average LRRi estimate for challenging behavior was 1.02, corresponding to an average change of -64%, 95% CI [-88%, +11%]. The average LRRi estimate for pro-social behavior was 0.83, corresponding to an average improvement of 129%, 95% [-67%, 1487%]. Coefficients of the potential moderators were all imprecisely estimated.