Online supplemental material

The following pages show the supplemental online material. In addition to this, the dataset as a whole will be provided in an Excel sheet.

Appendix

References for prior reviews that were searched

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S.M. (2015). Improving fluid intelligence with training on working memory: A meta-analysis. *Psychonomic Bulletin & Review*, 22, 366-377. DOI: 10.3758/s13423-014-0699-x

Conway, A. R. A., & Getz. S. J. (2010). Cognitive ability: Does working memory training enhance intelligence? *Current Biology, 20*, R362-R364. DOI: 10.1016/j.cub.2010.03.001

Cortese,S., Ferrin, M., Brandeis, D., Buitelaar, J., Daley, D., Dittmann, R.W., … Sonuga-Barke, E. (2015). Cognitive training for attention-deficit/hyperactivity disorder: Meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*(3), 164-174. DOI: 10.1016/j.jaac.2014.12.010

Hindin, S., & Zelinski, E. M. (2012). Extended practice and aerobic exercise interventions benefit untrained cognitive outcomes in older adults: A meta-analysis. *Journal of American Geriatric Society, 60*, 136–141. DOI: 10.1111/j.1532-5415.2011.03761.x

Karbach, J., & Verhaeghen, P. (2014). Making working memory work: A meta-analysis of executive-control and working memory training in older adults. *Psychological Science, 25*, 2027–2037. DOI: 10.1177/0956797614548725

Karch, D., Albers, L., Renner, G., Lichtenauer, N., & von Kries, R. (2013) The efficacy of cognitive training programs in children and adolescents: A meta-analysis. *Deutches Artzeblatt International, 110*, 643-652. DOI: 10.3238/arztebl.2013.0643

Karr, J. E., Areshenkoff, C. N., Rast, P., & Garcia-Barrera, M. A. (2014). An empirical comparison of the therapeutic benefits of physical exercise and cognitive training on the executive functions of older adults: A meta-analysis of controlled trials. *Neuropsychology, 28*, 1-17. DOI: 10.1037/neu0000101

Kelly, M. E., Loughrey, D., & Lawlor, B., (2014). The impact of cognitive training and mental stimulation on cognitive and everyday functioning of healthy older adults: A systematic review and meta-analysis. *Ageing Research Reviews*, *15*, 28-43. DOI: 10.1016/j.arr.2014.02.004

Morrison, A., & Chein, J. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review*, *18*, 46-60. DOI: 10.3758/s13423-010-0034-0

Peijnenborgh, J., Hurks, P. M., Aldenkamp, A. P., Vles, J.S., & Hendriksen, J.G. (2015). Efficacy of working memory training in children and adolescents with learning disabilities: A review study and meta-analysis. *Neuropsychological Rehabilitation: An International Journal*. DOI: 10.1080/09602011.2015.1026356

Schwaighofer, M., Fischer, F., & Buhner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist*, *50*(2), 138–166. DOI: 10.1080/00461520.2015.1036274

Shipstead, Z., Redick, T. S., & Engle, R.W. (2012). Is working memory training effective? *Psychological Bulletin, 138*(4), 628. DOI: 10.1037/a0027473

Sitzer, D. I., Twamley, E. W., & Jeste, D. V. (2006). Cognitive training in Alzheimer's disease: A meta-analysis of the literature. *Acta Psychiatrica Scandinavica, 114*(2), 75-90. DOI: 10.1111/j.1600-0447.2006.00789.x

Spencer-Smith, M., & Klingberg, T. (2015). Benefits of a working memory training program for inattention in daily life: A systematic review and meta-analysis. *PloS One*, *10*(3), e0119522. DOI: 10.1371/journal.pone.0119522

Toril, P., Reales, J. M., & Ballesteros, S. (2014). Video game training enhances cognition of older adults: A meta-analytic study. *Psychology and Aging*, *29*, 706–716. DOI: 10.1037/a0037507

Uttal, D. H., Meadow, N. G., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139* (2), 352-402. DOI: 10.1037/a0028446

von Bastian, C. C., & Oberauer, K. (2014). Effects and mechanisms of working memory training: A review. *Psychological Research*, *78*, 803-820. doi:10.1007/s00426-013-0524-6

Wass, S. V., Scerif, G., & Johnson, M. H. (2012). Training attentional control and working memory – Is younger, better? *Developmental Review*, 32, 360-387. DOI: 10.1016/j.dr.2012.07.001

Weicker, J., Villringer, A., & Thöne-Otto, A. (2015). Can impaired working memory functioning be improved by training? A meta-analysis with a special focus on brain injured patients. *Neuropsychology*. DOI: 10.1037/neu00002

Table s1

Characteristics of Working Memory Training Studies Included in the Meta-analysis

Study name	Comparison	Time point	Construct	Outcome test	Hedge's *g*	Sample size (t (c))	Age (t (c))
Alloway 2012	1	Posttest	Arithmetic	WOND	0.57	8 (7)	12.9 (13)
			Verbal WM	AWMA letters	1.51		
			Verbal reasoning	WASI Vocab	1.12		
Alloway et al 2013	1	Posttest	Arithmetic	WOND	-0.17	23 (32)	11.2 (10.6)
			Verbal reasoning	WASI Vocab	0.76		
			Verbal WM	AWMA composite	0.83		
			Visual WM	AWMA Shape recall	0.66		
	2	Posttest	Arithmetic	WOND	0.17	23 (39)	11.2 (10.1)
			Verbal reasoning	WASI	0.49		
			Verbal WM	AWMA composite	1.04		
			Visual WM	AWMA Shape recall	0.66		
	1	Follow up	Arithmetic	WOND	0.34	11 (19)	11.2 (10.6)
			Verbal reasoning	WASI	0.63		
			Verbal WM	AWMA composite	1.40		
			Visual WM	AWMA Shape recall	1.44		
	2	Follow up	Arithmetic	WOND	0.33	11 (24)	11.2 (10.11)
			Verbal reasoning	WASI	1.39		
			Verbal WM	AWMA composite	1.27		
			Visual WM	AWMA Shape recall	1.64		
Ang et al 2015	1	Posttest	Verbal WM 1	Animal updating	0.20	32 (28)	7 (7)
			Verbal WM 2	Back letter recall	0.42		
			Visual WM	Letter rotation	0.44		
			Arithmetic 1	Numerical Operations (WIAT-III)	0.05		
			Arithmetic 2	Addition fluency (WIAT-III)	0.25		
			Arithmetic 3	Subtraction fluency (WIAT-III)	0.49		

2	Posttest	Verbal WM 1	Animal updating	0.05	32 (26)	7 (7)
		Verbal WM 2	Back letter recall	0.44		
		Visual WM	Letter rotation	0.53		
		Arithmetic 1	Numerical Operations (WIAT-III)	0.16		
		Arithmetic 2	Addition fluency (WIAT-III)	0.24		
		Arithmetic 3	Subtraction fluency (WIAT-III)	0.08		
3	Posttest	Verbal WM 1	Animal updating	0.19	25 (28)	7 (7)
		Verbal WM 2	Back letter recall	0.67		
		Visual WM	Letter rotation	-0.09		
		Arithmetic 1	Numerical Operations (WIAT-III)	-0.29		
		Arithmetic 2	Addition fluency (WIAT-III)	0.00		
		Arithmetic 3	Subtraction fluency (WIAT-III)	-0.15		
4	Posttest	Verbal WM 1	Animal updating	0.04	25 (26)	7 (7)
		Verbal WM 2	Back letter recall	0.69		
		Visual WM	Letter rotation	-0.05		
		Arithmetic 1	Numerical Operations (WIAT-III)	-0.21		
		Arithmetic 2	Addition fluency (WIAT-III)	-0.04		
		Arithmetic 3	Subtraction fluency (WIAT-III)	-0.61		
1	Follow up	Verbal WM 1	Animal updating	0.43	32 (28)	7 (7)
		Verbal WM 2	Back letter recall	0.13		
		Visual WM	Letter rotation	0.34		
		Arithmetic 1	Numerical Operations (WIAT-III)	0.43		
		Arithmetic 2	Addition fluency (WIAT-III)	0.56		
		Arithmetic 3	Subtraction fluency (WIAT-III)	0.48		
2	Follow up	Verbal WM 1	Animal updating	-0.08	32 (26)	7 (7)
		Verbal WM 2	Back letter recall	0.14		
		Visual WM	Letter rotation	0.23		
		Arithmetic 1	Numerical Operations (WIAT-III)	-0.08		
		Arithmetic 2	Addition fluency (WIAT-III)	0.53		
		Arithmetic 3	Subtraction fluency (WIAT-III)	0.56		
3	Follow up	Verbal WM 1	Animal updating	0.37	25 (28)	7 (7)

			Verbal WM 2	Back letter recall	0.71		
			Visual WM	Letter rotation	-0.29		
			Arithmetic 1	Numerical Operations (WIAT-III)	0.11		
			Arithmetic 2	Addition fluency (WIAT-III)	-0.07		
			Arithmetic 3	Subtraction fluency (WIAT-III)	0.10		
	4	Follow up	Verbal WM 1	Animal updating	-0.16	25 (26)	7 (7)
			Verbal WM 2	Back letter recall	0.75		
			Visual WM	Letter rotation	-0.42		
			Arithmetic 1	Numer Operations (WIAT-III)	-0.38		
			Arithmetic 2	Addition fluency (WIAT-III)	-0.19		
			Arithmetic 3	Subtraction fluency (WIAT-III)	0.03		
Anguera et al 2012[1]	1	Posttest	Nonverbal reasoning 1	Card rotation	0.24	22 (22)	19 (19)
			Verbal WM	Operation span	0.54	22 (22)	
			Visual WM 1	4-back objects	0.84	22 (22)	
			Visual WM 2	3-back objects	0.64	22 (22)	
			Verbal reasoning 1	Analogies	0	29 (27)	
			Verbal reasoning 2	Letter Sets	-0.14	29 (27)	
			Nonverbal reasoning 2	Raven advanced	-0.01	29 (27)	
			Nonverbal reasoning 3	BOMAT	0.11	29 (27)	
Ashman-East 2015	1	Posttest	Verbal WM	AWMA listening recall	1.19	15 (13)	Fifth graders
			Visual WM	AWMA spatial recall	0.38		
			Arithmetic	Jamaica Grade 4 literacy test	0.56		
	1	Follow up	Arithmetic	Jamaica Grade 4 literacy test	0.53		
Bergman-Nutley & Klingberg 2014	1	Posttest	Visual WM	odd one out	0.74	155 (275)	11.1 (11.01)
			Arithmetic	math	0.19		
Borella et al 2014	1	Posttest	Crit	Matrix span	1.82	20 (20)	69.9 (69.55)
			Nonverbal reasoning	Cattell	0.29		
			Verbal WM	CWMS	1.94		
			Visual WM	Backward Corsi block	1.59		
	2	Posttest	Crit	Matrix span	1.35	20 (20)	79.6 (79.7)

			Nonverbal reasoning	Cattell	0.07		
			Verbal WM	CWMS	1.06		
			Visual WM	Backward Corsi block	-0.20		
	1	Follow up	Crit	Matrix span	1.41	20 (20)	69.9 (69.55)
			Nonverbal reasoning	Cattell	-0.07		
			Verbal WM	CWMS	2.01		
			Visual WM	Backward Corsi block	0.55		
	2	Follow up	Crit	Matrix span	0.96	20 (20)	79.6 (79.7)
			Nonverbal reasoning	Cattell	-0.09		
			Verbal WM	CWMS	1.13		
			Visual WM	Backward Corsi	-0.20		
Brehmer et al 2012	1	Posttest	Crit 1	Span board fw	1.74	29 (26)	26.2 (25.7)
			Crit 2	Span board bw	1.70		
			Crit 3	Digit span fw	0.34		
			Crit 4	Digit span bw	1.16		
			Nonverbal reasoning	Raven standard	-0.12		
	2	Posttest	Crit 1	Span board fw	0.81	26 (19)	63.9 (63.6)
			Crit 2	Span board bw	1.29		
			Crit 3	Digit span fw	0.56		
			Crit 4	Digit span bw	0.34		
			Nonverbal reasoning	Raven standard	-0.36		
	1	Follow up	Crit 1	Span board fw	2.00	29 (26)	26.2 (25.7)
			Crit 2	Span board bw	1.34		
			Crit 3	Digit span fw	0.32		
			Crit 4	Digit span bw	1.03		
			Nonverbal reasoning	Raven standard	-0.15		
	2	Follow up	Crit 1	Span board fw	0.94	26 (19)	63.9 (63.6)
			Crit 2	Span board bw	1.61		
			Crit 3	Digit span fw	0.67		
			Crit 4	Digit span bw	-0.08		
			Nonverbal reasoning	Raven standard	-0.28		

Bürki et al 2014	1	Posttest	Crit	Verbal 2 back	0.56	22 (20)	24.68 (24.35)
			Nonverbal reasoning	Raven	0.01		
			Verbal WM 1	Reading span	0.09		
			Verbal WM 2	Number updating	0.38		
			Visual WM	Spatial 2 back	0.33		
	2	Posttest	Crit	Verbal 2 back	0.60	22 (21)	24.68 (25.35)
			Nonverbal reasoning	Raven	-0.01		
			Verbal WM 1	Reading span	-0.14		
			Verbal WM 2	Number updating	0.22		
			Visual WM	Spatial 2 back	0.00		
	3	Posttest	Crit	Verbal 2 back	0.83	22 (20)	67.64 (67.7)
			Nonverbal reasoning	Raven	0.13		
			Verbal WM 1	Reading span	0.14		
			Verbal WM 2	Number updating	-0.37		
			Visual WM	Spatial 2 back	0.45		
	4	Posttest	Crit	Verbal 2 back	0.60	22 (23)	67.64 (68.61)
			Nonverbal reasoning	Raven	-0.09		
			Verbal WM 1	Reading span	0.02		
			Verbal WM 2	Number updating	-0.55		
			Visual WM	Spatial 2 back	0.44		
Chacko et al 2014	1	Posttest	Arithmetic	WRAT math computation	0.10	44 (41)	8.4 (8.4)
			Decoding 1	WRAT word reading	-0.05		
			Reading comp	WRAT sentence comprehension	0.31		
			Verbal WM	AWMA Listening recall	0.29		
			Visual WM	AWMA Spatial recall	0.07		
			Crit	AWMA Digit recall	0.28		
Chein & Morrison 2010[2]	1	Posttest	Crit	Verbal & Spatial CWM composite	1.39	22 (20)	20.1 (20.6)
			Nonverbal reasoning	Raven advanced	0.06		
			Reading comp	Nelson Denny reading comprehension	0.57	21 (19)	
Chooi & Thompson 2012[3]	1	Posttest	Nonverbal reasoning 1	ETS Card Rotation	0.24	9 (15)	~20 (~20)
			Nonverbal reasoning 2	ETS Paper Folding	0.71		

			Nonverbal reasoning 3	Mental rotation	-0.29		
			Nonverbal reasoning 4	Raven advanced	0.10		
			Verbal WM	Operation span	0.17		
			Verbal reasoning	Mill Hill vocab	0.53		
	2	Posttest	Nonverbal reasoning 1	ETS Card Rotation	0.43	9 (22)	~20 (~20)
			Nonverbal reasoning 2	ETS Paper Folding	0.45		
			Nonverbal reasoning 3	Mental rotation	-0.92		
			Nonverbal reasoning 4	Raven advanced	0.43		
			Verbal WM	Operation span	0.14		
			Verbal reasoning	Mill Hill vocab	0.24		
	3	Posttest	Nonverbal reasoning 1	ETS Card Rotation	-0.31	13 (11)	~20 (~20)
			Nonverbal reasoning 2	ETS Paper Folding	-0.26		
			Nonverbal reasoning 3	Mental rotation	0.11		
			Nonverbal reasoning 4	Raven advanced	0.26		
			Verbal WM	Operation span	0.12		
			Verbal reasoning	Mill Hill vocab	-0.15		
	4	Posttest	Nonverbal reasoning 1	ETS Card Rotation	-0.56	13 (23)	~20 (~20)
			Nonverbal reasoning 2	ETS Paper Folding	0.19		
			Nonverbal reasoning 3	Mental rotation	0.36		
			Nonverbal reasoning 4	Raven advanced	0.15		
			Verbal WM	Operation span	0.08		
			Verbal reasoning	Mill Hill vocab	-0.19		
Clouter 2013	1	Posttest	Nonverbal reasoning	Cattell	0.63	18 (18)	20.39 (20.39)
			Visual WM	Symmetry span	-0.02	17 (18)	
			Verbal WM	Operation span	0.57	18 (18)	
Colom et al 2013	1	Posttest	Nonverbal reasoning 1	Raven advanced	0.30	28 (28)	18.0 (18.2)
			Nonverbal reasoning 2	DAT-Abstract Reasoning	0.23		
			Verbal reasoning 1	PMA-R	0.05		
			Verbal reasoning 2	DAT-Verbal Reasoning	-0.31		
			Verbal reasoning 3	PMA-V	-0.19		
			Arithmetic	DAT-Number Reasoning	0.32		

			Verbal WM 1	Computation span	-0.26		
			Verbal WM 2	Reading span	0.14		
			Visual WM	Dot matrix	0.27		
Dahlin 2011 / 2013[4]	1	Posttest	Decoding 1	Word decoding	0.39	41 (15)	10.7 (10.7)
			Decoding 2	Orthographic verification	-0.24	29 (15)	
			Reading comp	PIRLS & IEA RSL narrative texts	0.56	41 (15)	
			Arithmetic 1	Addition	0.43	42 (15)	10.7 (10.7)
			Arithmetic 2	Subtraction	0.00	42 (15)	
			Arithmetic 3	Basic number screening test	0.24	42 (15)	
Dahlin et al 2008	1	Posttest	Crit	Letter working memory task	0.98	15 (11)	23.67 (24.09)
			Verbal WM 1	Digit span backward	-0.13		
			Verbal WM 2	3-back digit	0.50		
			Verbal WM 3	Computation span	-0.42		
			Nonverbal reasoning	Raven advanced	0.29		
	2	Posttest	Crit	Letter working memory task	1.12	13 (16)	68.38 (68.25)
			Verbal WM 1	Digit span backward	0.59		
			Verbal WM 2	3-back digit	-0.19		
			Verbal WM 3	Computation span	0.11		
			Nonverbal reasoning	Raven advanced	0.06		
	1	Follow up	Crit	Letter working memory task	1.01	11 (7)	23.67 (24.09)
			Verbal WM 1	Digit span backward	-0.01		
			Verbal WM 2	3-back digit	0.57		
			Verbal WM 3	Computation span	0.19		
			Nonverbal reasoning	Raven advanced	-0.14		
	2	Follow up	Crit	Letter working memory task	1.59	13 (7)	68.38 (68.25)
			Verbal WM 1	Digit span backward	0.25		
			Verbal WM 2	3-back digit	-0.15		
			Verbal WM 3	Computation span	0.17		
			Nonverbal reasoning	Raven advanced	0.28		
Dunning et al 2013[5]	1	Posttest	Arithmetic 1	WOND math reasoning	-0.10	33 (29)	8.42 (8.42)
			Arithmetic 2	WOND number operations	-0.21	33 (29)	

			Decoding 1	WORD basic reading	0.22	34 (17)	
			Decoding 2	NARA reading accuracy	0.01	30 (28)	
			Decoding 3	NARA reading rate	0.03	29 (27)	
			Nonverbal reasoning	WASI Performance IQ composite	-0.36	34 (27)	
			Reading comp	NARA reading comprehension	0.21	30 (28)	
			Verbal reasoning	WASI Verbal IQ composite	0.16	34 (26)	
			Verbal WM	AWMA verbal composite	1.56	34 (30)	
			Visual WM	AWMA visuo-spatial composite	1.05	34 (30)	
	2	Posttest	Arithmetic 1	WOND math reasoning	-0.15	33 (30)	9.25 (9.5)
			Arithmetic 2	WOND number operations	-0.13	33 (30)	
			Decoding 1	WORD basic reading	-0.23	34 (30)	
			Decoding 2	NARA reading accuracy	0.04	30 (28)	
			Decoding 3	NARA reading rate	0.23	29 (27)	
			Nonverbal reasoning	WASI Performance IQ composite	-0.30	24 (30)	
			Reading comp	NARA reading comprehension	0.22	30 (28)	
			Verbal reasoning	WASI Verbal IQ composite	-0.14	24 (30)	
			Verbal WM	AWMA composite	2.24	34 (30)	
			Visual WM	AWMA composite	1.07	34 (30)	
	1	Follow up	Arithmetic 1	WOND math reasoning	-0.13	15 (16)	
			Decoding 1	WORD basic reading	-0.06	14 (17)	
			Decoding 2	NARA reading accuracy	-0.18	14 (17)	
			Decoding 3	NARA reading rate	-0.67	14 (17)	
			Nonverbal reasoning	WASI matrix reasoning	-0.22	14 (17)	
			Reading comp	NARA reading comprehension	-0.09	14 (17)	
			Verbal reasoning	WASI similarities	-0.10	14 (17)	
			Verbal WM	AWMA backward digit recall	1.30	15 (19)	
			Visual WM	AWMA Mr X	0.21	15 (19)	
Egeland et al 2013 / Hovik et al 2013[6]	1	Posttest	Arithmetic	Key math	0.27	33 (34)	10.5 (10.3)
			Decoding 1	LOGOS decoding quality	0.55		
			Decoding 2	LOGOS decoding speed	-0.32		

			Verbal WM 1	Digit span fw and bw	0.59		
			Visual WM	Leiter span fw and bw	0.66		
			Verbal WM 2	Let-Num Sequence and Sentence span	0.72		
	1	Follow up	Arithmetic	Key math	0.23	33 (34)	10.5 (10.3)
			Decoding 1	LOGOS decoding quality	0.62		
			Decoding 2	LOGOS decoding speed	-0.15		
			Verbal WM 1	Digit span fw and bw	0.47		
			Visual WM	Leiter span fw and bw	1.13		
			Verbal WM 2	Let-Num Sequence and Sentence span	0.25		
Estrada et al 2015	1	Posttest	Nonverbal reasoning 1	Raven advanced	-0.01	170 (193)	20.3 (20.3)
			Nonverbal reasoning 2	DAT-Abstract Reasoning	-0.04		
			Nonverbal reasoning 3	DAT-Spatial Reasoning	-0.14		
			Verbal reasoning	DAT-Verbal Reasoning	0.04		
Everts et al 2015	1	Posttest	Verbal WM	WISC-IV Letter-Number Sequence	0.40	22(23)	9.45 (9.34)
			Visual WM	BASIC MLT Spatial positioning	0.15		
			Nonverbal reasoning	WISC-IV Matrices	-0.07		
			Decoding	ELFE Sentence Reading	-0.26		
			Arithmetic	WISC-IV Arithmetic	0.04		
	1	Follow up	Verbal WM	WISC-IV Letter-Number Sequence	0.27	22(23)	
			Visual WM	BASIC MLT Spatial positioning	0.02		
			Nonverbal reasoning	WISC-IV Matrices	-0.19		
			Decoding	ELFE Sentence Reading	-0.43		
			Arithmetic	WISC-IV Arithmetic	-0.36		
Feiyue et al 2009	1	Posttest	Nonverbal reasoning	Raven standard	1.54	8 (10)	24.25 (24.15)
Foster et al 2014	1	Posttest	Visual WM 1	Reading span with images	0.28	20 (18)	18-35 (18-35)
			Visual WM 2	Rotation span	0.46	20 (18)	
			Visual WM 3	Running shape span	0.01	20 (18)	
			Visual WM 4	Running icons span	-0.11	20 (18)	
			Visual WM 5	Change detection orientation	-0.12	20 (18)	
			Verbal WM	Keep Track	0.11	20 (18)	
			Nonverbal reasoning 1	Matrix reasoning	-0.11	20 (18)	

			Nonverbal reasoning 2	Paper Folding	-0.21	20 (18)	
	2	Posttest	Visual WM 1	Reading span with images	0.70	20 (19)	18-35 (18-35)
			Visual WM 2	Rotation span	1.73	20 (19)	
			Visual WM 3	Running shape span	0.60	20 (19)	
			Visual WM 4	Running icons span	0.16	20 (19)	
			Visual WM 5	Change detection orientation	-0.28	20 (19)	
			Verbal WM	Keep Track	-0.01	20 (19)	
			Nonverbal reasoning 1	Matrix reasoning	-0.11	20 (19)	
			Nonverbal reasoning 2	Paper Folding	0.05	20 (19)	
	3	Posttest	Visual WM 1	Reading span with images	0.10	19 (18)	18-35 (18-35)
			Visual WM 2	Rotation span	-0.29	19 (18)	
			Visual WM 3	Running shape span	-0.33	19 (18)	
			Visual WM 4	Running icons span	-0.07	19 (18)	
			Visual WM 5	Change detection orientation	0.12	19 (18)	
			Verbal WM	Keep Track	0.07	19 (18)	
			Nonverbal reasoning 1	Matrix reasoning	0.22	19 (18)	
			Nonverbal reasoning 2	Paper Folding	-0.05	19 (18)	
	4	Posttest	Visual WM 1	Reading span with images	0.51	20 (19)	18-35 (18-35)
			Visual WM 2	Rotation span	-0.20	20 (19)	
			Visual WM 3	Running shape span	0.07	20 (19)	
			Visual WM 4	Running icons span	-0.17	20 (19)	
			Visual WM 5	Change detection orientation	0.09	20 (19)	
			Verbal WM	Keep Track	-0.15	20 (19)	
			Nonverbal reasoning 1	Matrix reasoning	-0.19	20 (19)	
			Nonverbal reasoning 2	Paper Folding	0.10	20 (19)	
Gray et al 2012[7]	1	Posttest	Decoding 1	WRAT word reading	0.04	31 (21)	14.2 (14.4)
			Reading comp	WRAT sentence comprehension	-0.09		
			Verbal WM	WISC digit span backward	0.82		
			Visual WM 1	CANTAB spatial span fw and bw	-0.05		
			Visual WM 2	CANTAB spatial WM errors	0.17		
Gropper et al 2014	1	Posttest	Arithmetic 1	WAIS-IV arithmetic	0.15	39 (23)	28-30 (28-30)

			Arithmetic 2	W-J applied problems	-0.06		
			Crit 1	CANTAB spatial span	0.73		
			Crit 2	WAIS-IV digit span	0.33		
			Reading comp	Nelson Denny reading comprehension	-0.14		
			Verbal WM 1	PASAT a	-0.27		
			Verbal WM 2	PASAT b	0.12		
			Visual WM 1	CANTAB spatial WM errors	-0.20		
			Visual WM 2	CANTAB spatial strategy	0.23		
	1	Follow up	Arithmetic 1	WAIS-IV arithmetic	0.29	24 (21)	28-30 (28-30)
			Arithmetic 2	W-J applied problems	-0.07		
			Crit 1	CANTAB spatial span	1.46		
			Crit 2	WAIS-IV digit span	0.59		
			Reading comp	Nelson Denny reading comprehension	-0.24		
			Verbal WM 1	PASAT a	-0.22		
			Verbal WM 2	PASAT b	0.13		
			Visual WM 1	CANTAB spatial WM errors	-0.05		
			Visual WM 2	CANTAB spatial strategy	0.57		
Hanson 2013	1	Posttest	Crit 2	AWMA digit recall	-0.09	17(9)	12.35 (11.14)
			Crit 1	AWMA dot matrix	0.69		
			Verbal WM 1	AWMA listening recall	-0.01		
			Visual WM	AWMA spatial recall	-0.24		
			Verbal WM 2	WISC-IV Digit span/Let num seq	0.55		
			Reading comp 1	GORT comprehension	0.29		
			Reading comp 2	W-J comprehension	-0.19		
Harrison et al 2013	1	Posttest	Crit 1	Running letter span	0.82	21 (17)	undergrads
			Crit 2	Running matrix span	1.15		
			Nonverbal reasoning	Raven advanced	0.19		
			Verbal reasoning 1	Letter sets	0.03		
			Verbal reasoning 2	Number series	-0.13		
			Verbal WM 1	Reading span	1.08		
			Verbal WM 2	Keep track	0.73		

			Visual WM 1	Rotation span	1.39		
			Visual WM 2	Change detection	-0.18	21 (16)	
	2	Posttest	Crit 1	Running letter span	0.96	17 (17)	undergrads
			Crit 2	Running matrix span	0.96		
			Nonverbal reasoning	Raven advanced	0.24		
			Verbal reasoning 1	Letter sets	-0.12		
			Verbal reasoning 2	Number series	-0.49		
			Verbal WM 1	Reading span	0.31		
			Verbal WM 2	Keep track	0.89		
			Visual WM 1	Rotation span	0.18		
			Visual WM 2	Change detection	-0.09	17 (16)	
Heffernan 2014	1	Posttest	Nonverbal reasoning	Cattell	0.18	9(10)	22.1 (21.7)
			Verbal WM	Operation span	0.25		
			Visual WM	Symmetry span	0.01		
			Visual WM 1	Move span	0.57		
Heinzel et al 2014	1	Posttest	Verbal WM	Digit span backward	-0.06	15 (15)	25.9 (25.6)
			Nonverbal reasoning 1	LPS Figural Relations	0.28	14 (15)	
			Nonverbal reasoning 2	Raven standard	0.59		
	2	Posttest	Verbal WM	Digit span backward	0.60	15 (15)	66.07 (65.6)
			Nonverbal reasoning 1	LPS Figural Relations	0.33		
			Nonverbal reasoning 2	Raven standard	0.18		
Holmes et al 2009 [8]	1	Posttest	Arithmetic	WOND mathematical reasoning	-0.11	22 (20)	10.08 (9.75)
			Verbal WM	AWMA verbal WM composite	2.39		
			Decoding	WORD basic reading	-0.09		
			Nonverbal reasoning	WASI performance IQ	-0.19		
			Verbal reasoning	WASI verbal IQ	0.29		
			Visual WM	AWMA visuospatial WM	0.85		
Horvat 2014	1	Posttest	nonverbal reasoning	Raven standard	0.62	14(15)	14.07 (13.67)
Jaeggi et al 2008[9]	1	Posttest	Nonverbal reasoning	Raven advanced	0.59	8 (7)	24.48 (24.48)
			Verbal WM	Digit span backward	0.48		
	2	Posttest	Nonverbal reasoning	Raven advanced	-0.05	8 (8)	24.48 (24.48)

			Verbal WM	Digit span backward	0.31		
	3	Posttest	Nonverbal reasoning	BOMAT	0.62	11 (11)	undergrads
	4	Posttest	Nonverbal reasoning	BOMAT	0.88	8 (8)	26.1 (27.8)
	5	Posttest	Nonverbal reasoning	BOMAT	0.88	7 (8)	25.5 (25.13)
			Verbal WM 1	Digit span backward	2.02		
			Verbal WM 2	Reading span	1.05		
			Visual WM	Visuospatial span backward	-0.47		
Jaeggi et al 2011	1	Posttest	Nonverbal reasoning 1	Raven standard	0.13	32 (30)	9.25 (8.83)
			Nonverbal reasoning 2	TONI	0.00		
	1	Follow up	Nonverbal reasoning 1	Raven standard	-0.04		
			Nonverbal reasoning 2	TONI	-0.04		
Jaeggi et al 2014	1	Posttest	Nonverbal reasoning 1	BOMAT	0.16	25 (27)	25.21 (25.21)
			Nonverbal reasoning 2	Cattell	0.20		
			Nonverbal reasoning 3	Raven advanced	-0.07		
			Nonverbal reasoning 4	ETS Surface development	0.21		
			Nonverbal reasoning 5	ETS Form board	0.54		
			Nonverbal reasoning 6	ETS Space relations	0.17		
			Reading comp	AFOQT Reading comp	0.10		
			Verbal reasoning 1	AFOQT Verbal analogies	0.11	24 (27)	
			Verbal reasoning 2	ETS Inferences	-0.11		
	2	Posttest	Nonverbal reasoning 1	BOMAT	0.55	26 (27)	25.21(25.21)
			Nonverbal reasoning 2	Cattell	0.20		
			Nonverbal reasoning 3	Raven advanced	0.16		
			Nonverbal reasoning 4	ETS Surface development	0.08		
			Nonverbal reasoning 5	ETS Form board	0.29		
			Nonverbal reasoning 6	ETS Space relations	0.39		
			Reading comp	AFOQT Reading comp	0.37		
			Verbal reasoning 1	AFOQT Verbal analogies	0.38		
			Verbal reasoning 2	ETS Inferences	-0.37		
	1	Follow up	Nonverbal reasoning 1	BOMAT	0.26	17(23)	25.21 (22.79)
			Nonverbal reasoning 2	Cattell	0.27		

	2	Follow up	Nonverbal reasoning 1	BOMAT	0.19	14 (23)	25.21 (22.79)
			Nonverbal reasoning 2	Cattell	0.03		
Jaeggi et al 2010	1	Posttest	Visual WM	N-back object	1.32	25 (41)	19.1 (19.4)
			Nonverbal reasoning 1	BOMAT	0.29	25 (43)	
			Nonverbal reasoning 2	Raven advanced	0.78	25 (43)	
			Verbal WM	Operation span	-0.27	25 (40)	
	2	Posttest	Visual WM	N-back object	1.12	20 (41)	19.0 (19.4)
			Nonverbal reasoning 1	BOMAT	0.54	21 (43)	
			Nonverbal reasoning 2	Raven advanced	0.50	21 (43)	
			Verbal WM	Operation span	-0.41	21 (40)	
Jaušovec & Jaušovec 2012	1	Posttest	Nonverbal reasoning 1	Paper folding	0.26	14 (15)	20.25 (20.25)
			Verbal reasoning	Word analogies	0.29		
			Nonverbal reasoning 2	Raven standard/advanced	0.54		
			Verbal WM	Digit span backward	0.38		
Karbach et al 2014[10]	1	Posttest	Crit	Farm/Safari span	0.78	14 (14)	8.4 (8.4)
			Visual WM	Color span	0.65		
			Arithmetic	German Mathematics test	-0.26		
	1	Follow up	Visual WM	Color span	1.21	14 (12)	8.4 (8.4)
			Arithmetic	German Mathematics test	0.07		
Klingberg et al 2002	1	Posttest	Crit	Visuospatial WM	3.18	7 (7)	11.0 (11.4)
			Nonverbal reasoning	Raven colored	2.18		
			Visual WM	Span board forward and backward	1.66		
Klingberg et al 2005	1	Posttest	Nonverbal reasoning	Raven colored	0.23	20 (24)	9.75 (9.67)
			Visual WM	Span board forward and backward	0.77		
		Follow up	Nonverbal reasoning	Raven colored	0.05	18 (24)	9.75 (9.67)
			Visual WM	Span board forward and backward	0.79		
Kundu et al 2013[11]	1	Posttest	Nonverbal reasoning	Raven advanced	-0.01	10 (10)	20.9 (20.9)
			Verbal WM	Operation span	-0.36	13 (13)	
			Visual WM	Color-in-location task	0.22	13 (15)	
	2	Posttest	Nonverbal reasoning	Raven advanced (full)	0.00	3 (3)	20.9 (20.9)
Lange & Süß 2015	1	Posttest	Crit 1	Operation span	0.53	31 (31)	66.85 (68.23)

			Crit 2	Dot span	0.76		
			Crit 3	Memory updating numerical	0.90		
			Crit 4	Running figural span	1.16		
			Verbal WM 1	Reading span	0.01		
			Verbal WM 2	Swaps	0.00		
			Verbal WM 3	Running numerical span	0.32		
			Verbal reasoning 1	BIS-4 verbal reasoning	-0.06		
			Verbal reasoning 2	BIS-4 numerical reasoning	-0.21		
			Nonverbal reasoning	BIS-4 figural analogies	-0.15		
	2	Posttest	Crit 1	Operation span	0.33	31 (29)	66.85 (68.69)
			Crit 2	Dot span	0.87		
			Crit 3	Memory updating numerical	0.55		
			Crit 4	Running figural span	0.91		
			Verbal WM 1	Reading span	0.28		
			Verbal WM 2	Swaps	0.11		
			Verbal WM 3	Running numerical span	0.78		
			Verbal reasoning 1	BIS-4 verbal reasoning	0.14		
			Verbal reasoning 2	BIS-4 numerical reasoning	-0.46		
			Nonverbal reasoning	BIS-4 figural analogies	0.31		
Lee 2014[12]	1	Posttest	Verbal WM	Digit span back	0.21	25 (25)	9.2 (9.2)
			Visual WM	Visual 2-back correct	-0.39		
			Decoding 1	R-CBM	0.36		
			Decoding 2	W-J III Reading fluency	-0.11		
			Decoding 3	Maze-CBM	0.58		
			Reading comp 1	W-J III Passage comp	0.33		
			Reading comp 2	Gates-MacGinitie	0.13		
			Nonverbal reasoning	TONI	-0.03		
Lindeløv et al 2014	1	Posttest	Verbal WM 1	Wechsler WM index	0.15	9 (9)	29.2 (29.4)
			Nonverbal reasoning	Raven advanced	0.17		
			Verbal WM 2	Operation Span	0.02		
	2	Posttest	Verbal WM 1	Wechsler WM index	0.29	8 (9)	56.1 (56.1)

			Nonverbal reasoning	Raven advanced	-0.24		
			Verbal WM 2	Operation Span	0.01		
Loosli et al 2012[13]	1	Posttest	Decoding 1	Salzburger Lesetest psuedowords	-0.06	20 (20)	9.97 (10.02)
			Decoding 2	Salzburger Lesetest words	0.28		
			Decoding 3	Salzburger Lesetest text	0.39		
			Nonverbal reasoning	TONI	0.12		
Mansur-Alves & Flores-Mendoza 2015	1	Posttest	Nonverbal reasoning	Raven standard	0.07	27 (26)	11.17 (11.17)
			Arithmetic	BPR5 numerical reasoning	-0.12		
Mansur-Alves et al 2013[14]	1	Posttest	Nonverbal reasoning 1	Raven standard	0.06	8 (8)	8.75 (8.75)
			Nonverbal reasoning 2	TNVRI	0.37		
Minear et al 2012	1	Posttest	Arithmetic 1	ETS Mathematics aptitude	0.09	31 (26)	19.9 (19.8)
			Arithmetic 2	ETS Arithmetic aptitude	0.00		
			Nonverbal reasoning 1	Raven standard	-0.51		
			Nonverbal reasoning 2	Cattell	0.06		
			Reading comp	LSAT reading comp	-0.16		
			Verbal reasoning 1	Inferences	-0.05		
			Verbal reasoning 2	Nonsense syllogisms	0.05		
			Verbal WM 1	Letter n-back	1.06		
			Verbal WM 2	Operation span	0.30		
			Verbal WM 3	Letter-number span	-0.07		
			Verbal WM 4	Reading span	-0.23		
			Visual WM 1	Object n-back	0.52		
			Visual WM 2	Symmetry span	0.13		
			Visual WM 3	Rotation span	0.53		
			Visual WM 4	Alignment span	-0.17		
	2	Posttest	Arithmetic 1	ETS Mathematics aptitude	0.06	32 (26)	19.7 (19.8)
			Arithmetic 2	ETS Arithmetic aptitude	0.35		
			Nonverbal reasoning 1	Raven standard	-0.35		
			Nonverbal reasoning 2	Cattell	-0.30		
			Reading comp	LSAT reading comp	0.01		

			Verbal reasoning 1	Inferences	0.17		
			Verbal reasoning 2	Nonsense syllogisms	0.31		
			Verbal WM 1	Letter n-back	0.12		
			Verbal WM 2	Operation span	0.46		
			Verbal WM 3	Letter-number span	-0.29		
			Verbal WM 4	Reading span	-0.13		
			Visual WM 1	Object n-back	0.06		
			Visual WM 2	Symmetry span	-0.12		
			Visual WM 3	Rotation span	0.73		
			Visual WM 4	Alignment span	0.05		
	3	Posttest	Arithmetic 1	ETS Mathematics aptitude	0.02	27 (26)	19.6 (19.8)
			Arithmetic 2	ETS Arithmetic aptitude	0.47		
			Nonverbal reasoning 1	Raven standard	-0.23		
			Nonverbal reasoning 2	Cattell	-0.23		
			Reading comp	LSAT reading comp	-0.02		
			Verbal reasoning 1	Nonsense syllogisms	0.26		
			Verbal reasoning 2	Inferences	0.05		
			Verbal WM 1	Letter n-back	0.75		
			Verbal WM 2	Operation span	0.27		
			Verbal WM 3	Letter-number span	-0.30		
			Verbal WM 4	Reading span	-0.04		
			Visual WM 1	Object n-back	0.61		
			Visual WM 2	Symmetry span	0.03		
			Visual WM 3	Rotation span	0.73		
			Visual WM 4	Alignment span	0.08		
Minear et al 2013	1	Posttest	Nonverbal reasoning	Raven advanced	0.04	42 (23)	undergrads
			Nonverbal reasoning	Raven standard	0.06		
			Verbal WM	Reading span	0.16		
Moreau et al 2015	1	Posttest	Nonverbal reasoning 1	Surface development	-0.08	21 (22)	29.73 (29.73)
			Nonverbal reasoning 2	Form board	0.04		
			Nonverbal reasoning 3	Mental rotation	0.19		

			Nonverbal reasoning 4	Paper folding	0.16		
			Verbal WM 1	Backward digit span	0.36		
			Verbal WM 2	Letter number sequencing	0.49		
Nussbaumer et al 2013[15]	1	Posttest	Arithmetic 1	Mental arithmetics	0.39	29 (27)	23.7 (23.7)
			Arithmetic 2	Mathematik-test	-0.23		
			Nonverbal reasoning	Raven advanced	0.38		
			Verbal WM	Operation span	-0.07		
	2	Posttest	Arithmetic 1	Mental arithmetics	0.51	27 (27)	23.7 (23.7)
			Arithmetic 2	Mathematik-test	0.18		
			Nonverbal reasoning	Raven advanced	0.48		
			Verbal WM	Operation span	-0.21		
	1	Follow up	Arithmetic	Mathematik-test	0.43	29 (27)	23.7 (23.7)
			Nonverbal reasoning	Raven advanced	0.13		
			Verbal WM	Operation span	-0.15		
	2	Follow up	Arithmetic	Mathematik-test	0.40	27 (27)	23.7 (23.7)
			Nonverbal reasoning	Raven advanced	0.00		
			Verbal WM	Operation span	0.00		
Nutley et al 2011	1	Posttest	Verbal WM	Odd one out	0.89	24 (25)	4.27 (4.27)
			Nonverbal reasoning 1	Raven colored Set A	-0.10		
			Nonverbal reasoning 2	Raven colored Set AB	-0.10		
			Nonverbal reasoning 3	Raven colored Set B	-0.56		
			Nonverbal reasoning 4	WPPSI Block design	0.11		
Oelhafen et al 2013	1	Posttest	Crit	Dual n-back	3.27	14 (15)	25.2 (25.2)
			Nonverbal reasoning	BOMAT	-0.05		
			Verbal WM	Reading span	-0.11		
	2	Posttest	Crit	Dual n-back	0.20	14 (15)	25.2 (25.2)
			Nonverbal reasoning	BOMAT	0.00		
			Verbal WM	Reading span	0.03		
Payne 2014	1	Posttest	Crit	Reading span	0.59	22(19)	67.68 (68.11)
			Verbal WM 1	Listening span	1.05		
			Verbal WM 2	Operation span	0.72		

			Verbal WM 3	Minus 2 span	0.38		
			Reading comp	Nelson Denny reading comprehension	0.00		
Pugin et al 2015	1	Posttest	Verbal WM	Auditory letter n-back	0.73	14 (15)	10-16
			Nonverbal reasoning	TONI	0.37		
			Verbal WM	Letter number sequencing	0.18		
		Follow up	Verbal WM	Auditory letter n-back	1.13	14 (15)	10-16
			Nonverbal reasoning	TONI	0.63		
			Verbal WM	Letter number sequencing	0.21		
Redick et al 2013	1	Posttest	Nonverbal reasoning 1	Raven standard	0.00	24 (29)	21.1 (20.7)
			Nonverbal reasoning 2	Raven advanced	-0.05		
			Nonverbal reasoning 3	Cattell	-0.43		
			Nonverbal reasoning 4	Paper folding	0.30		
			Verbal reasoning 1	General knowledge	-0.08		
			Verbal reasoning 2	Verbal analogies	-0.13		
			Verbal reasoning 3	Inferences	0.31		
			Verbal reasoning 4	Number series	-0.12		
			Verbal WM	Running letter span	0.11		
			Visual WM	Symmetry span	0.34		
	2	Posttest	Nonverbal reasoning 1	Raven standard	0.04	24 (20)	21.1 (21.2)
			Nonverbal reasoning 2	Raven advanced	-0.18		
			Nonverbal reasoning 3	Cattell	-0.05		
			Nonverbal reasoning 4	Paper folding	0.36		
			Verbal reasoning 1	General knowledge	0.30		
			Verbal reasoning 2	Verbal analogies	0.20		
			Verbal reasoning 3	Inferences	0.16		
			Verbal reasoning 4	Number series	0.32		
			Verbal WM	Running Letter span	-0.02		
			Visual WM	Symmetry span	0.26		
Redick & Wiemers 2015	1	Posttest	Verbal WM 1	Running letter span	0.42	30 (29)	20.5 (20.5)
			Visual WM 1	Running matrix span	-0.20		
			Verbal WM 2	Letter 3-back	0.36		

			Visual WM 2	Matrix 3-back	-0.15		
			Visual WM 3	Color change detection	-0.28		
			Visual WM 4	Orient change detection	0.14		
			Reading comp	Nelson Denny reading comprehension	0.18		
	2	Posttest	Verbal WM 1	Running letter span	-0.09	27 (29)	20.4 (20.5)
			Visual WM 1	Running matrix span	-0.28		
			Verbal WM 2	Letter 3-back	0.17		
			Visual WM 2	Matrix 3-back	0.04		
			Visual WM 3	Color change detection	-0.06		
			Visual WM 4	Orient change detection	0.18		
			Reading comp	Nelson Denny reading comprehension	0.34		
Reimer et al 2014	1	Posttest	Visual WM 1	WRAML finger windows	0.52	16 (13)	24 (24)
			Verbal WM 1	WRAML number letters	0.35	16 (13)	
			Verbal WM 2	WRAML verbal working memory	0.57	15 (12)	
			Verbal WM 3	WRAML symbolic working memory	0.18	16 (13)	
			Verbal WM 4	Operation span	0.19	16 (13)	
			Visual WM 2	Symmetry span	0.17	16 (12)	
			Nonverbal reasoning	Raven advanced	0.09	16 (13)	
			Verbal reasoning 1	Letter sets	0.25	16 (13)	
			Verbal reasoning 2	Inferences	-0.26	16 (13)	
Richey et al 2014[16]	1	Posttest	Visual WM	Spatial WM	0.58	25 (24)	18-30
			Nonverbal reasoning	Raven advanced	0.28	25 (24)	
	2	Posttest	Visual WM	Spatial WM	0.39	25 (26)	18-30
			Verbal reasoning	Analogies	-0.19	25 (24)	
Richmond et al 2011	1	Posttest	Nonverbal reasoning	Raven	-0.40	21 (19)	66
			Verbal WM 1	Digit span backward	-0.50		
			Verbal WM 2	Reading span	0.67		
Rode et al 2014[17]	1	Posttest	Arithmetic 1	WIAT-II mathematical reasoning	0.04	156 (126)	8-9
			Reading comp	WIAT-II reading comprehension	0.11		
			Arithmetic 2	CMB math	0.14		
			Decoding	CMB reading fluency	-0.05		

Rudebeck et al 2012	1	Posttest	Crit	Dual n-back	3.81	27 (28)	25.36 (25.49)
			Nonverbal reasoning	BOMAT	0.71		
Salminen et al 2012	1	Posttest	Crit	Dual n-back	5.61	20 (16)	24.4 (24.5)
			Nonverbal reasoning	Raven advanced	-0.68	13 (9)	
Savage 2013	1	Posttest	Nonverbal reasoning 1	Cattell	0.23	23 (27)	46.65 (48.44)
			Nonverbal reasoning 2	Raven advanced	-0.11		
			Visual WM	Spatial manipulation	-0.52		
			Verbal WM 1	WAIS-IV Digit span fw, bw, and seq	-0.09		
			Verbal WM 2	Operation span	0.08	23 (26)	
Schwarb et al 2015 E1[18]	1	Posttest	Nonverbal reasoning	Raven advanced	-0.12	27 (25)	18-30
Schwarb et al 2015 E2	1	Posttest	Nonverbal reasoning	Cattell	0.35	22 (22)	18-32
			Nonverbal reasoning	Raven advanced	0.07		
			Verbal WM	Operation span	0.45	21 (21)	
			Visual WM 1	Symmetry span	0.13		
			Visual WM 2	Change detection	1.04		
	2	Posttest	Nonverbal reasoning	Cattell	-0.06	22 (22)	18-32
			Nonverbal reasoning	Raven advanced	0.50		
			Verbal WM	Operation span	0.47	22 (21)	
			Visual WM 1	Symmetry span	0.63		
			Visual WM 2	Change detection	0.83		
Schweizer et al 2011	1	Posttest	Nonverbal reasoning	Raven standard	1.37	14 (16)	25 (25)
	2	Posttest	Nonverbal reasoning	Raven standard	0.89	15 (16)	25 (25)
Shavelson et al 2008	1	Posttest	Nonverbal reasoning	Raven	0.01	18 (19)	13.5 (13.5)
			Verbal WM 1	Reading span	0.15		
			Verbal WM 2	Operation Span	0.33		
			Crit	Span board forward & backward	0.52		
			Verbal WM 3	Digit span forward & backward	0.97		
Shiran & Breznitz 2011[19]	1	Posttest	Crit 1	Cognifit visuo-spatial WM	0.36	26 (15)	24.84
			Crit 2	Cognifit auditory WM	0.67		
			Crit 3	Cognifit visual-verbal WM	0.39		
			Decoding 1	Words per minute	0.46		

			Decoding 2	Pseudowords per minute	0.31		
			Reading comp	Silent reading comprehension	0.87		
			Decoding 3	Parsing test (number correct)	-0.08		
	2	Posttest	Crit 1	Cognifit visuo-spatial WM	0.50	35 (15)	25.11
			Crit 2	Cognifit auditory WM	0.26		
			Crit 3	Cognifit visual-verbal WM	0.50		
			Decoding 1	Words per minute	0.29		
			Decoding 2	Pseudowords per minute	0.52		
			Reading comp	Silent reading comprehension	0.58		
			Decoding 3	Parsing test (number correct)	-0.03		
Smith et al 2013[20]	1	Posttest	Nonverbal reasoning	Raven advanced	0.04	10 (10)	18-34
	2	Posttest	Nonverbal reasoning	Raven advanced	-0.64	10 (9)	18-34
	3	Posttest	Nonverbal reasoning	Raven advanced	0.00	10 (10)	18-34
Söderqvist et al 2012[21]	1	Posttest	Nonverbal reasoning	Raven colored	-0.01	22 (19)	9.68
			Nonverbal reasoning	WPPSI Block design	0.25		
			Verbal WM	Word span backward	0.38		
			Visual WM	Odd one out	0.39		
	1	Follow up	Nonverbal reasoning	Raven colored	-0.14	22 (19)	9.68
			Nonverbal reasoning	WPPSI Block design	-0.15		
			Verbal WM	Word span backward	-0.07		
			Visual WM	Odd one out	-0.01		
Sprenger et al 2013 E1[22]	1	Posttest	Verbal WM 1	Operation span	0.49	58 (55)	22.97 (23.05)
			Verbal WM 2	Listening span	0.28	58 (54)	
			Visual WM 1	Symmetry span	0.34	59 (55)	
			Visual WM 2	Rotation span	0.14	55 (52)	
			Reading comp	AFOQT Reading comprehension	-0.16	59 (54)	
			Verbal reasoning 1	ETS Inferences	-0.25	59 (54)	
			Verbal reasoning 2	AFOQT Verbal analogies	0.07	59 (55)	
	1	Follow up	Verbal WM 1	Operation span	0.69	49 (47)	
			Verbal WM 2	Listening span	0.34	48 (46)	
			Visual WM 1	Symmetry span	0.60	49 (47)	

			Visual WM 2	Rotation span	0.42	49 (44)	
			Reading comp	AFOQT Reading comprehension	-0.10	46 (46)	
			Verbal reasoning 1	ETS Inferences	-0.20	46 (46)	
			Verbal reasoning 2	AFOQT Verbal analogies	0.01	46 (46)	
Sprenger et al 2013 E2	1	Posttest	Nonverbal reasoning	Raven advanced	0.15	33 (37)	35.51
			Reading comp	Nelson Denny reading comprehension	0.06	34 (37)	
			Verbal WM	Reading span	-0.20	33 (44)	
			Visual WM	Shapebuilder	1.07	30 (35)	
			Crit	Letter N-back	0.67	34 (37)	
	2	Posttest	Crit	Shapebuilder	1.40	34 (37)	35.51
			Nonverbal reasoning	Raven advanced	0.13	34 (37)	
			Reading comp	Nelson Denny reading comprehension	0.12	33 (37)	
			Verbal WM 1	Reading span	-0.22	34 (37)	
			Verbal WM 2	Letter N-back	0.00	33 (37)	
	3	Posttest	Crit 1	Shapebuilder	1.07	34 (37)	35.51
			Nonverbal reasoning	Raven advanced	0.12	33 (37)	
			Reading comp	Nelson Denny reading comprehension	-0.05	34 (37)	
			Verbal WM	Reading span	-0.37	34 (37)	
			Crit 2	Letter N-back	0.74	33 (37)	
Stepankova et al 2014	1	Posttest	Crit	Letter N-back	2.29	20 (25)	67.95 (68.08)
			Nonverbal reasoning 1	WASI Block design	0.12		
			Nonverbal reasoning 2	WASI Matrix reasoning	0.18		
			Verbal WM 1	WMS-III Digit span fw & bw	0.17		
			Verbal WM 2	WMS-III Letter-Number Sequencing	0.67		
	2	Posttest	Crit	Letter N-back	1.96	20 (25)	68.15 (68.08)
			Nonverbal reasoning 1	WASI Block design	0.42		
			Nonverbal reasoning 2	WASI Matrix reasoning	0.49		
			Verbal WM 1	WMS-III Digit span fw & bw	0.16		
			Verbal WM 2	WMS-III Letter-Number Sequencing	0.51		
Stephenson & Halpern 2013	1	Posttest	Nonverbal reasoning 1	Beta-III Matrix reasoning	1.18	28 (26)	22.48
			Nonverbal reasoning 2	Cattell	0.18		

			Nonverbal reasoning 3	Paper folding	0.52		
			Nonverbal reasoning 4	Mental rotation	-0.03		
			Nonverbal reasoning 5	Raven advanced	0.71		
			Nonverbal reasoning 6	WASI Matrix reasoning	0.40		
			Verbal reasoning	Extended range vocabulary	0.08		
	2	Posttest	Nonverbal reasoning	Beta-III Matrix reasoning	0.68	29 (26)	22.48
			Nonverbal reasoning	Cattell	0.18		
			Nonverbal reasoning	Paper folding	0.42		
			Nonverbal reasoning	Mental rotation	-0.10		
			Nonverbal reasoning	Raven advanced	0.46		
			Nonverbal reasoning	WASI Matrix reasoning	0.37		
			Verbal reasoning	Extended range vocabulary	0.00		
	3	Posttest	Nonverbal reasoning	Beta-III Matrix reasoning	0.37	25 (26)	22.48
			Nonverbal reasoning	Cattell	0.06		
			Nonverbal reasoning	Paper folding	0.18		
			Nonverbal reasoning	Mental rotation	0.17		
			Nonverbal reasoning	Raven advanced	0.39		
			Nonverbal reasoning	WASI Matrix reasoning	0.25		
			Verbal reasoning	Extended range vocabulary	-0.04		
	4	Posttest	Nonverbal reasoning	Beta-III Matrix reasoning	0.58	28 (26)	22.48
			Nonverbal reasoning	Cattell	-0.07		
			Nonverbal reasoning	Paper folding	0.33		
			Nonverbal reasoning	Mental rotation	0.09		
			Nonverbal reasoning	Raven advanced	0.57		
			Nonverbal reasoning	WASI Matrix reasoning	0.45		
			Verbal reasoning	Extended range vocabulary	0.02		
Takeuchi et al 2013	1	Posttest	Arithmetic 1	Simple arithmetic	0.04	41 (20)	20.9 (21.4)
			Arithmetic 2	Complex arithmetic	0.05		
			Arithmetic 3	Kyodai SX test	0.09		
			Nonverbal reasoning 1	BOMAT	-0.31		
			Nonverbal reasoning 2	Raven advanced	0.46		

			Verbal WM	Digit span fw & bw	1.29		
			Visual WM	Visuospatial WM fw & bw	0.50		
Thompson et al 2013	1	Posttest	Crit	Dual N-back	2.77	20 (19)	21.2 (23.1)
			Decoding	Nelson Denny reading rate	0.24	20 (18)	
			Nonverbal reasoning 1	Raven advanced	0.23	20 (19)	
			Nonverbal reasoning 2	WASI/WAIS Matrix reasoning	-0.37	20 (19)	
			Nonverbal reasoning 3	WASI/WAIS Blocks	-0.10	20 (19)	
			Reading comp	Nelson Denny reading comprehension	0.00	20 (18)	
			Verbal WM 1	Operation span	0.18	19 (14)	
			Verbal WM 2	Reading span	0.21	20 (14)	
			Verbal reasoning 1	WASI/WAIS Similarities	0.24	20 (19)	
			Verbal reasoning 2	WASI/WAIS Vocabulary	-0.05	20 (19)	
	2	Posttest	Crit	Dual N-back	2.67	20 (19)	21.2 (21.3)
			Decoding	Nelson Denny reading rate	-0.04		
			Nonverbal reasoning 1	Raven advanced	0.10		
			Nonverbal reasoning 2	WASI/WAIS Matrix reasoning	-0.81		
			Nonverbal reasoning 3	WASI/WAIS Blocks	-0.22		
			Reading comp	Nelson Denny reading comprehension	0.00		
			Verbal WM 1	Operation span	0.46	19 (19)	
			Verbal WM 2	Reading span	0.47		
			Verbal reasoning 1	WASI/WAIS Similarities	0.21		
			Verbal reasoning 2	WASI/WAIS Vocabulary	0.00		
Thorell et al 2009	1	Posttest	Nonverbal reasoning	WISC Block design	-0.03	17 (14)	4.5 (4.8)
			Verbal WM	Word span fw & bw	1.09		
			Visual WM	WAIS-R-NI Span board fw & bw	0.45		
	2	Posttest	Nonverbal reasoning	WISC Block design	0.33	17 (16)	4.5 (5)
			Verbal WM	Word span fw & bw	1.06		
			Visual WM	WAIS-R-NI Span board fw & bw	0.70		
Urbánek & Marček 2015	1	Posttest	Nonverbal reasoning 1	Raven advanced	0.33	31 (34)	24.8 (24.6)
			Nonverbal reasoning 2	BOMAT	0.01		
	2	Posttest	Nonverbal reasoning 1	Raven advanced	0.04	37 (34)	25.7 (24.6)

			Nonverbal reasoning 2	BOMAT	-0.25		
Van der Molen et al 2010	1	Posttest	Arithmetic	De Vos 1992 arithmetic	0.00	41 (26)	15.32 (15.43)
			Decoding	Brus & Voeten 1973 reading	0.09		
			Nonverbal reasoning	Raven standard	-0.23		
			Verbal WM 1	WMTB-C Backward digit recall	0.22		
			Verbal WM 2	WMTB-C Listening recall	0.09		
			Visual WM	AWMA Spatial span	0.17		
	2	Posttest	Arithmetic	De Vos 1992 arithmetic test	-0.05	26 (26)	15 (15.43)
			Decoding	Brus & Voeten 1973 reading test	0.06		
			Nonverbal reasoning	Raven standard	0.03		
			Verbal WM 1	WMTB-C Backward digit recall	-0.15		
			Verbal WM 2	WMTB-C Listening recall	0.04		
			Visual WM	AWMA Spatial span	0.14		
	1	Follow up	Arithmetic	De Vos 1992 arithmetic	0.10	39 (25)	15 (15.43)
			Decoding	Brus & Voeten 1973 reading	0.17		
			Nonverbal reasoning	Raven standard	-0.23		
			Verbal WM 1	WMTB-C Backward digit recall	0.20		
			Verbal WM 2	WMTB-C Listening recall	0.06		
			Visual WM	AWMA Spatial span	0.42		
	2	Follow up	Arithmetic	De Vos 1992 arithmetic test	-0.07	25 (25)	15 (15.43)
			Decoding	Brus & Voeten 1973 reading test	0.04		
			Nonverbal reasoning	Raven standard	-0.07		
			Verbal WM 1	WMTB-C Backward digit recall	0.06		
			Verbal WM 2	WMTB-C Listening recall	0.06		
			Visual WM	AWMA Spatial span	0.03		
van Dongen-Boomsma et al 2014	1	Posttest	Nonverbal reasoning	Raven colored	-0.08	26 (21)	6.5 (6.6)
			Verbal WM	Digit span WISC-III backward	0.91	22 (21)	
			Visual WM	Knox cubes LDT backward	0.12	26 (19)	
Vartanian et al 2013	1	Posttest	Nonverbal reasoning	Raven advanced	0.63	17 (17)	30.79
von Bastian & Eschen 2015	1	Posttest	Crit 1	Numerical span	0.95	34 (32)	23 (23)

			Crit 2	Verbal span	0.66		
			Crit 3	Figural span	1.16		
			Nonverbal reasoning 1	Locations test	-0.31		
			Nonverbal reasoning 2	Diagramming relationships	-0.30		
			Nonverbal reasoning 3	Raven advanced	0.02		
			Verbal reasoning	Nonsense syllogisms	0.04		
			Verbal WM 1	Binding	0.06		
			Verbal WM 2	Brown-peterson	0.38		
			Verbal WM 3	Number updating	0.16	34 (31)	
	2	Posttest	Crit 1	Numerical span	0.96	30 (32)	22.5 (23)
			Crit 2	Verbal span	0.66		
			Crit 3	Figural span	1.04		
			Nonverbal reasoning 1	Locations test	-0.15		
			Nonverbal reasoning 2	Diagramming relationships	0.14		
			Nonverbal reasoning 3	Raven advanced	-0.15		
			Verbal reasoning	Nonsense syllogisms	-0.07		
			Verbal WM 1	Binding	-0.36		
			Verbal WM 2	Brown-peterson	0.12		
			Verbal WM 3	Number updating	0.26		
	3	Posttest	Crit 1	Numerical span	0.96	34 (32)	23.12 (23)
			Crit 2	Verbal span	0.55		
			Crit 3	Figural span	0.90		
			Nonverbal reasoning 1	Locations test	-0.32		
			Nonverbal reasoning 2	Diagramming relationships	0.04		
			Nonverbal reasoning 3	Raven advanced	-0.28		
			Verbal reasoning	Nonsense syllogisms	-0.23		
			Verbal WM 1	Binding	0.13		
			Verbal WM 2	Brown-peterson	0.06		
			Verbal WM 3	Number updating	0.21	34 (31)	23 (23)
von Bastian & Oberauer 2013[23]	1	Posttest	Verbal reasoning	Syllogisms	-0.08	30 (30)	22.87 (23.77)

			Verbal WM	Memory updating	0.07		
	1	Follow up	Verbal reasoning	Syllogisms	0.24	30 (30)	22.87 (23.77)
			Verbal WM	Memory updating	-0.07		
Wang et al 2014	1	Posttest	Nonverbal reasoning	Raven standard	0.54	20 (20)	10-11
	2	Posttest	Nonverbal reasoning	Raven standard	0.25	20 (20)	
	3	Posttest	Nonverbal reasoning	Raven standard	0.20	20 (20)	
	4	Posttest	Nonverbal reasoning	Raven standard	0.01	15 (20)	
Weicker et al 2013	1	Posttest	Verbal WM	Digit span bw	0.00	10 (5)	67.3 (67.0)
			Visual WM 1	Span board bw	1.11	12 (7)	
			Visual WM 2	Symbol span	-0.47	12 (7)	
			Nonverbal reasoning	LPS Figural Relations	0.09	12 (7)	
	2	Posttest	Verbal WM	Digit span bw	-0.62	10 (13)	67.3 (67.6)
			Visual WM 1	Span board bw	0.77	12 (13)	
			Visual WM 2	Symbol span	-0.62	12 (13)	
			Nonverbal reasoning	LPS Figural Relations	0.14	12 (13)	
Westerberg et al 2007	1	Posttest	Nonverbal reasoning	Raven standard	-0.10	9 (9)	55 (53.6)
Xin et al 2014	1	Posttest	Verbal WM 1	Numerical updating	0.84	15 (14)	70 (69)
			Nonverbal reasoning	Raven advanced	0.09		
			Verbal WM 2	WAIS-R Digit span bw	1.15		
Zhang et al 2014	1	Posttest	Nonverbal reasoning 1	BOMAT	0.65	26 (24)	23
		Posttest	Nonverbal reasoning 2	Form Board	0.40		
		Posttest	Nonverbal reasoning 3	Space Relations	0.29		
		Posttest	Nonverbal reasoning 4	ETS Surface Development	0.39		
Zhao et al 2011	1	Posttest	Nonverbal reasoning	Raven standard	0.54	16 (17)	9.76
Zinke et al 2014	1	Posttest	Nonverbal reasoning	Raven standard	0.39	40 (40)	76.7 (77.7)
			Verbal WM	Letter span plus	0.85		
		Follow up	Nonverbal reasoning	Raven standard	0.04	33 (18)	76.7 (77.7)
			Verbal WM	Letter span plus	0.62		

Note. The Comparison column provides a label for a specific comparison within that study; studies with multiple comparisons have multiple labels; t = training, c = control, WM = working memory, crit = criterion near transfer measure, Reading comp = reading comprehension

[1]Anguera et al 2012 – Analogies, Letter Sets, Raven, BOMAT not reported in published article
[2]Chein & Morrison 2010 – ETS Reasoning battery not coded because it was a composite of nonverbal and verbal reasoning outcomes
[3]Chooi & Thompson 2012 – Additional vocabulary outcome not coded because of administration issues reported in article
[4]Dahlin 2011/2013 – Same subjects, different outcomes reported. Also, WM and Raven outcomes not coded given the control group did not perform those tasks (control data reported in article for those outcomes are from Klingberg et al, 2005). Also, follow-up data not included because as stated in Dahlin (2013, p. 123), part of the training group completed an additional 10 training sessions during the posttest-to-follow-up interval, and their data were combined with the training participants that did not complete any additional training.
[5]Dunning et al 2013 – Untreated control group did not complete follow-up outcomes
[6]Egeland et al 2013/Hovik et al 2013 – Same subjects, different outcomes reported
[7]Gray et al 2012 – WRAT math not coded because the active control group performed arithmetic tasks, producing a negative effect size because the control group improved more from pretest to posttest than the WM training group
[8]Holmes et al 2009 – Follow-up outcomes not coded because the control group did not complete follow-up outcomes
[9]Jaeggi et al 2008 – Active control group data not reported in published article
[10]Karbach et al 2014 – Knuspels reading ability test not coded because it was a composite of reading comprehension and decoding outcomes
[11]Kundu et al 2013 – Location VSTM not coded as TMS pulses were applied on 50% of trials
[12]Lee 2014 – TONI not reported in dissertation
[13]Loosli et al 2012 – Reading time used as dependent variable instead of errors given that errors were so infrequent
[14]Mansur-Alves et al 2013 – *Bateria Fatorial CEPA* and *Teste de Desempnho Escolar* not coded because each test was a composite of multiple outcomes (verbal reasoning, arithmetic, decoding)
[15]Nussbaumer et al 2013 – *Intelligenz-Struktur-Test* not coded because it was a composite of multiple outcomes (verbal reasoning, nonverbal reasoning, arithmetic)
[16]Richey et al 2014 – Active control group data and Raven outcome not reported in published article (presented in Phillips et al. 2012 conference poster)
[17]Rode et al 2014 – AWMA composite not coded because it was a composite of multiple outcomes (verbal WM, visual WM)
[18]Schwarb et al 2015 E1 – pretest and posttest means and SDs only available for Raven advanced progressive matrices outcome
[19]Shiran & Breznitz 2011 – Sternberg memory task not coded because was administered with EEG recording for training groups only
[20]Smith et al 2013 – Follow-up data not coded because the follow-up session was only 1 week after the posttest, in contrast to all other studies with follow-up comparisons that occurred months after posttest.
[21]Söderqvist et al 2012 – Aston Index test and Allet Teller test not coded because relevant data not available in text, and not assessed at posttest; authors did report that "training had no effect on outcome measures employed in this study assessing cognitive abilites or school assessments at the T3 follow-up" (p. 5).
[22]Sprenger et al 2013 E1 – Inconsistent sample sizes due to attrition, and for reading comprehension and verbal reasoning follow-up

outcomes, authors provided pretest and posttest means/SDs for only those subjects that completed study

[23]von Bastian & Oberauer 2013 – BIS-4s reasoning not coded because it was a composite of digit series, figure series, figural analogies, word analogies, fact/opinion, and estimation subtests; Brown-Peterson not coded because it was a composite of verbal WM and visual WM

Table s2

Categorization of Moderators in Working Memory Training Studies Included in Meta-Analysis

Study	Comp	Training Type	Training Content	Dose	Control Type	Random	Age	Learner Status	Publication
Alloway 2012	1	Adaptive Other	Both	Large	Passive	Yes	Child	Atypical	Published
Alloway et al 2013	1	Adaptive Other	Both	Large	Active	Yes	Child	Atypical	Published
Alloway et al 2013	2	Adaptive Other	Both	Large	Passive	Yes	Child	Atypical	Published
Ang et al 2015	1	Adaptive Running	Visual	Large	Active	No	Child	Atypical	Published
Ang et al 2015	2	Adaptive Running	Visual	Large	Passive	No	Child	Atypical	Published
Ang et al 2015	3	Adaptive Cogmed	Visual	Large	Active	No	Child	Atypical	Published
Ang et al 2015	4	Adaptive Cogmed	Visual	Large	Passive	No	Child	Atypical	Published
Anguera et al 2012[1]	1	Adaptive N-back *	Both	Small	Active	Yes	Y. Adult	Typical	Published
Ashman-East 2015	1	Adaptive Cogmed	Both	Large	Active	Yes	Child	Atypical	Grey
Bergman-Nutley & Klingberg 2014[2]	1	Adaptive Cogmed	Both	Large	Passive	No	Child	Atypical	Published
Borella et al 2014	1	Adaptive Complex	Visual	Small	Passive	Yes	O. Adult	Typical	Published
Borella et al 2014	2	Adaptive Complex	Visual	Small	Passive	Yes	O. Adult	Typical	Published
Brehmer et al 2012	1	Adaptive	Both	Large	Active	Yes	Y. Adult	Typical	Published

Brehmer et al 2012	2	Cogmed Adaptive Cogmed	Both	Large	Active	Yes	O. Adult	Typical	Published
Bürki et al 2014	1	Adaptive N-back	Verbal	Small	Active	No	Y. Adult	Typical	Published
Bürki et al 2014	2	Adaptive N-back	Verbal	Small	Passive	No	Y. Adult	Typical	Published
Bürki et al 2014	3	Adaptive N-back	Verbal	Small	Active	No	O. Adult	Typical	Published
Bürki et al 2014	4	Adaptive N-back	Verbal	Small	Passive	No	O. Adult	Typical	Published
Chacko et al 2014	1	Adaptive Cogmed	Both	Large	Active	Yes	Child	Atypical	Published
Chein & Morrison 2010	1	Adaptive Complex	Both	Large	Passive	Yes	Y. Adult	Typical	Published
Chooi & Thompson 2012	1	Adaptive N-back *	Both	Small	Active	No	Y. Adult	Typical	Published
Chooi & Thompson 2012	2	Adaptive N-back *	Both	Small	Passive	No	Y. Adult	Typical	Published
Chooi & Thompson 2012	3	Adaptive N-back *	Both	Large	Active	No	Y. Adult	Typical	Published
Chooi & Thompson 2012	4	Adaptive N-back *	Both	Large	Passive	No	Y. Adult	Typical	Published
Clouter 2013	1	Adaptive N-back *	Both	Small	Active	Yes	Y. Adult	Typical	Grey
Colom et al 2013	1	Adaptive N-back *	Both	Large	Passive	No	Y. Adult	Typical	Published
Dahlin 2011/2013	1	Adaptive Cogmed	Both	Large	Passive	No	Child	Atypical	Published
Dahlin et al 2008	1	Adaptive Other	Both	Large	Passive	Yes	Y. Adult	Typical	Published
Dahlin et al 2008	2	Adaptive Other	Both	Large	Passive	Yes	O. Adult	Typical	Published

Dunning et al 2013	1	Adaptive Cogmed	Both	Large	Active	No	Child	Atypical	Published
Dunning et al 2013	2	Adaptive Cogmed	Both	Large	Passive	No	Child	Atypical	Published
Egeland et al 2013/Hovik et al 2013	1	Adaptive Cogmed	Both	Large	Passive	Yes	Child	Atypical	Published
Estrada et al 2015	1	Non-Adaptive Other	Both	Small	Passive	Yes	Y. Adult	Typical	Published
Everts et al 2015	1	Adaptive Other	Both	Small	Passive	No	Child	Atypical	Published
Feiyue et al 2009	1	Adaptive N-back *	Both	Small	Passive	No	Y. Adult	Typical	Published
Foster et al 2014	1	Adaptive Complex	Both	Large	Active	Yes	Y. Adult	Atypical	Grey
Foster et al 2014	2	Adaptive Complex	Both	Large	Active	Yes	Y. Adult	Typical	Grey
Foster et al 2014	3	Adaptive Other	Both	Large	Active	Yes	Y. Adult	Atypical	Grey
Foster et al 2014	4	Adaptive Other	Both	Large	Active	Yes	Y. Adult	Typical	Grey
Gray et al 2012	1	Adaptive Cogmed	Both	Large	Active	Yes	Child	Atypical	Published
Gropper et al 2014	1	Adaptive Cogmed	Both	Large	Passive	Yes	Y. Adult	Atypical	Published
Hanson 2013	1	Adaptive Cogmed	Both	Large	Active	Yes	Child	Atypical	Grey
Harrison et al 2013	1	Adaptive Complex	Both	Large	Active	Yes	Y. Adult	Typical	Published
Harrison et al 2013	2	Adaptive Other	Both	Large	Active	Yes	Y. Adult	Typical	Published
Heffernan 2014	1	Adaptive N-back *	Both	Small	Active	Yes	Y. Adult	Typical	Grey
Heinzel et al 2014	1	Adaptive	Verbal	Small	Passive	Yes	Y. Adult	Typical	Published

Heinzel et al 2014	2	N-back Adaptive N-back	Verbal	Small	Passive	Yes	O. Adult	Typical	Published
Holmes et al 2009	1	Adaptive Cogmed	Both	Large	Active	No	Child	Atypical	Published
Horvat 2014	1	Adaptive N-back *	Both	Small	Passive	No	Child	Typical	Grey
Jaeggi et al 2008[3]	1	Adaptive N-back *	Both	Small	Active	No	Y. Adult	Typical	Published
Jaeggi et al 2008	2	Adaptive N-back *	Both	Small	Passive	No	Y. Adult	Typical	Published
Jaeggi et al 2008	3	Adaptive N-back *	Both	Small	Passive	No	Y. Adult	Typical	Published
Jaeggi et al 2008	4	Adaptive N-back *	Both	Small	Passive	No	Y. Adult	Typical	Published
Jaeggi et al 2008[3]	5	Adaptive N-back *	Both	Small	Passive	No	Y. Adult	Typical	Published
Jaeggi et al 2011	1	Adaptive N-back	Visual	Small	Active	No	Child	Typical	Published
Jaeggi et al 2014	1	Adaptive N-back *	Both	Small	Active	No	Y. Adult	Typical	Published
Jaeggi et al 2014	2	Adaptive N-back	Verbal	Small	Active	No	Y. Adult	Typical	Published
Jaeggi et al 2010	1	Adaptive N-back *	Both	Small	Passive	No	Y. Adult	Typical	Published
Jaeggi et al 2010	2	Adaptive N-back	Visual	Small	Passive	No	Y. Adult	Typical	Published
Jaušovec & Jaušovec 2012[4]	1	Adaptive Other	Both	Large	Active	No	Y. Adult	Typical	Published
Karbach et al 2014	1	Adaptive Complex	Visual	Small	Active	Yes	Child	Typical	Published
Klingberg et al 2002[5]	1	Adaptive Cogmed	Both	Large	Passive	Yes	Child	Atypical	Published

Klingberg et al 2005	1	Adaptive Cogmed	Both	Large	Active	Yes	Child	Atypical	Published
Kundu et al 2013	1	Adaptive N-back *	Both	Large	Active	Yes	Y. Adult	Typical	Published
Kundu et al 2013[6]	2	Adaptive N-back *	Both	Large	Active	Yes	Y. Adult	Typical	Published
Lange & Süß 2015[7]	1	Adaptive Other	Both	Large	Active	Yes	O. Adult	Typical	Published
Lange & Süß 2015[7]	2	Adaptive Other	Both	Large	Passive	Yes	O. Adult	Typical	Published
Lee 2014	1	Adaptive Complex	Visual	Small	Active	Yes	Child	Typical	Grey
Lindeløv et al 2014	1	Adaptive N-back	Both	Small	Active	No	Y. Adult	Typical	Grey
Lindeløv et al 2014	2	Adaptive N-back	Both	Small	Active	No	Y. Adult	Atypical	Grey
Loosli et al 2012	1	Adaptive Complex	Visual	Small	Passive	No	Child	Typical	Published
Mansur-Alves & Flores-Mendoza 2015	1	Adaptive Other	Verbal	Large	Passive	Yes	Child	Typical	Published
Mansur-Alves et al 2013	1	Adaptive Other	Verbal	Large	Active	Yes	Child	Typical	Published
Minear et al 2012	1	Adaptive N-back	Visual	Small	Active	No	Y. Adult	Typical	Grey
Minear et al 2012	2	Adaptive Complex	Verbal	Large	Active	No	Y. Adult	Typical	Grey
Minear et al 2012	3	Non-Adaptive N-back	Visual	Small	Active	No	Y. Adult	Typical	Grey
Minear et al 2013[8]	1	Adaptive Other	Both	Large	Passive	Yes	Y. Adult	Typical	Grey
Moreau et al 2015	1	Adaptive Complex	Both	Large	Active	Yes	Y. Adult	Typical	Published
Nussbaumer et al 2013	1	Adaptive	Both	Small	Active	Yes	Y. Adult	Typical	Published

Nussbaumer et al 2013	2	N-back * Non-Adaptive Other	Both	Small	Active	Yes	Y. Adult	Typical	Published
Nutley et al 2011	1	Adaptive Cogmed	Both	Small	Active	Yes	Child	Typical	Published
Oelhafen et al 2013	1	Adaptive N-back *	Both	Small	Passive	Yes	Y. Adult	Typical	Published
Oelhafen et al 2013	2	Adaptive N-back *	Both	Small	Active	Yes	Y. Adult	Typical	Published
Payne 2014	1	Adaptive Complex	Verbal	Small	Active	Yes	O. Adult	Typical	Grey
Pugin et al 2015	1	Adaptive N-back	Visual	Small	Passive	No	Child	Typical	Published
Redick et al 2013	1	Adaptive N-back *	Both	Large	Active	No	Y. Adult	Typical	Published
Redick et al 2013	2	Adaptive N-back *	Both	Large	Passive	No	Y. Adult	Typical	Published
Redick & Wiemers 2015	1	Adaptive Complex	Verbal	Small	Active	Yes	Y. Adult	Typical	Grey
Redick & Wiemers 2015	2	Adaptive Complex	Verbal	Small	Active	Yes	Y. Adult	Typical	Grey
Reimer et al 2014	1	Adaptive Other	Both	Large	Active	Yes	Y. Adult	Typical	Grey
Richey et al 2014[9]	1	Adaptive Complex	Both	Small	Active	No	Y. Adult	Typical	Published
Richey et al 2014	2	Adaptive Complex	Both	Small	Passive	No	Y. Adult	Typical	Published
Richmond et al 2011	1	Adaptive Complex	Both	Small	Active	Yes	O. Adult	Typical	Published
Rode et al 2014	1	Adaptive Complex	Verbal	Small	Passive	No	Child	Typical	Published
Rudebeck et al 2012	1	Adaptive N-back	Visual	Small	Passive	Yes	Y. Adult	Typical	Published

Salminen et al 2012	1	Adaptive N-back *	Both	Large	Passive	Yes	Y. Adult	Typical	Published
Savage 2013	1	Adaptive N-back *	Both	Small	Active	Yes	Y. Adult	Typical	Grey
Schwarb et al 2015 E1	1	Adaptive N-back	Both	Small	Passive	Yes	Y. Adult	Typical	Published
Schwarb et al 2015 E2	1	Adaptive N-back	Visual	Small	Passive	Yes	Y. Adult	Typical	Published
Schwarb et al 2015 E2	2	Adaptive N-back	Verbal	Small	Passive	Yes	Y. Adult	Typical	Published
Schweizer et al 2011	1	Adaptive N-back *	Both	Small	Active	Yes	Y. Adult	Typical	Published
Schweizer et al 2011	2	Adaptive N-back *	Both	Small	Active	Yes	Y. Adult	Typical	Published
Shavelson et al 2008	1	Adaptive Cogmed	Both	Large	Active	Yes	Child	Atypical	Published
Shiran & Breznitz 2011	1	Adaptive Other	Both	Large	Active	No	Y. Adult	Atypical	Published
Shiran & Breznitz 2011	2	Adaptive Other	Both	Large	Active	No	Y. Adult	Typical	Published
Smith et al 2013	1	Adaptive N-back *	Both	Small	Active	Yes	Y. Adult	Typical	Published
Smith et al 2013	2	Adaptive N-back *	Both	Small	Passive	Yes	Y. Adult	Typical	Published
Smith et al 2013	3	Adaptive N-back *	Both	Small	Active	Yes	Y. Adult	Typical	Published
Söderqvist et al 2012[10]	1	Adaptive Other	Visual	Small	Active	No	Child	Atypical	Published
Sprenger et al 2013 E1	1	Adaptive Other	Both	Large	Passive	Yes	Y. Adult	Typical	Published
Sprenger et al 2013 E2[11]	1	Adaptive Other	Verbal	Large	Active	Yes	Y. Adult	Typical	Published
Sprenger et al 2013 E2	2	Adaptive	Visual	Large	Active	Yes	Y. Adult	Typical	Published

		Other							
Sprenger et al 2013 E2[10]	3	Adaptive Other	Both	Large	Active	Yes	Y. Adult	Typical	Published
Stepankova et al 2014	1	Adaptive N-back	Verbal	Small	Passive	Yes	O. Adult	Typical	Published
Stepankova et al 2014	2	Adaptive N-back	Verbal	Small	Passive	Yes	O. Adult	Typical	Published
Stephenson & Halpern 2013	1	Adaptive N-back *	Both	Small	Passive	Yes	Y. Adult	Typical	Published
Stephenson & Halpern 2013	2	Adaptive N-back	Visual	Small	Passive	Yes	Y. Adult	Typical	Published
Stephenson & Halpern 2013	3	Adaptive N-back	Verbal	Small	Passive	Yes	Y. Adult	Typical	Published
Stephenson & Halpern 2013	4	Adaptive Other	Visual	Small	Passive	Yes	Y. Adult	Typical	Published
Takeuchi et al 2013	1	Adaptive Other	Both	Large	Passive	No	Y. Adult	Typical	Published
Thompson et al 2013	1	Adaptive N-back *	Both	Large	Passive	No	Y. Adult	Typical	Published
Thompson et al 2013	2	Adaptive N-back *	Both	Large	Active	No	Y. Adult	Typical	Published
Thorell et al 2009	1	Adaptive Cogmed	Visual	Small	Active	No	Child	Typical	Published
Thorell et al 2009	2	Adaptive Cogmed	Visual	Small	Passive	No	Child	Typical	Published
Urbánek & Marček 2015	1	Adaptive N-back	Both	Small	Active	Yes	Y. Adult	Typical	Published
Urbánek & Marček 2015	2	Adaptive N-back	Visual	Small	Active	Yes	Y. Adult	Typical	Published
Van der Molen et al 2010	1	Adaptive Complex	Visual	Small	Active	Yes	Child	Atypical	Published
Van der Molen et al 2010	2	Adaptive Complex	Visual	Small	Active	Yes	Child	Atypical	Published

van Dongen-Boomsma et al 2014	1	Adaptive Cogmed	Visual	Small	Active	Yes	Child	Atypical	Published
Vartanian et al 2013	1	Non-Adaptive N-back	Verbal	Small	Active	Yes	Y. Adult	Typical	Published
von Bastian & Eschen 2015	1	Adaptive Complex	Both	Large	Active	Yes	Y. Adult	Typical	Published
von Bastian & Eschen 2015	2	Random Complex	Both	Large	Active	Yes	Y. Adult	Typical	Published
von Bastian & Eschen 2015	3	Self-selected Complex	Both	Large	Active	Yes	Y. Adult	Typical	Published
von Bastian & Oberauer 2013	1	Adaptive Complex	Both	Large	Active	Yes	Y. Adult	Typical	Published
Wang et al 2014	1	Adaptive Other	Both	Small	Passive	Yes	Child	Typical	Published
Wang et al 2014	2	Adaptive Other	Both	Small	Passive	Yes	Child	Typical	Published
Wang et al 2014	3	Adaptive Other	Both	Small	Passive	Yes	Child	Typical	Published
Wang et al 2014	4	Adaptive Other	Both	Small	Passive	Yes	Child	Typical	Published
Weicker et al 2013	1	Adaptive Other	Both	Small	Active	Yes	O. Adult	Typical	Grey
Weicker et al 2013	2	Adaptive Other	Both	Small	Passive	Yes	O. Adult	Typical	Grey
Westerberg et al 2007	1	Adaptive Cogmed	Both	Large	Passive	Yes	Y. Adult	Atypical	Published
Xin et al 2014	1	Adaptive Other	Both	Small	Active	No	O. Adult	Typical	Published
Zhang et al 2014	1	Adaptive N-back	Both	Small	Active	No	Y. Adult	Typical	Grey
Zhao et al 2011	1	Adaptive Other	Visual	Small	Active	No	Child	Typical	Published
Zinke et al 2014	1	Adaptive	Both	Small	Passive	Yes	O. Adult	Typical	Published

Other	

Note. The Comparison column provides a label for a specific comparison within that study; studies with multiple comparisons have multiple labels; asterisk denotes dual N-back training type.

[1]Anguera et al 2012 – coded publication status as 'published', although four of the nonverbal ability outcomes were not included in the published article
[2]Bergman-Nutley & Klingberg 2014 – coded learner status as 'atypical', although control group was composed of typically developing children
[3]Jaeggi et al 2008 – coded publication status as 'published', although active-control group data in comparison 1 and two working memory outcomes in comparison 5 were not included in the published article
[4]Jaušovec & Jaušovec 2012 –2 of the 5 training tasks were not computerized
[5]Klingberg et al 2002 – 80% of training was WM (Cogmed); note that control group was coded as 'passive' because control participants completed only 10 trials (compared to 30 trials in training group), but coding the study as 'active' does not change the significance any of the outcomes.
[6]Kundu et al 2013 – comparison 2 was subset of n = 3 participants in each group that completed full version of Raven at pretest and posttest; all other participants (comparison 1) completed half at pretest and other half at posttest
[7]Lange & Süß 2015 – 4 of the 5 training tasks were WM (80%)
[8]Minear et al 2013 – 50% of training was WM
[9]Richey et al 2014 – coded publication status as 'published', although the nonverbal ability outcome was not included in the published article
[10]Söderqvist et al 2012 – 50% of training was WM (Cogmed)
[11]Sprenger et al 2013 E2 – for comparison 1, 50% of training was WM (N-back); for comparison 4, 75% of training was WM

Table s3

Number of participants (and mean sample size) of studies for each outcome measure

Construct	Time point	Total number of participants training group (mean sample size)	Total number of participants control group (mean sample size)
Near transfer measure	Posttest	975 (23.8)	912 (22.2)
	Follow up	143 (20.4)	120 (17.1)
Verbal working memory	Posttest	2118 (23.3)	2053 (22.6)
	Follow up	556 (25.3)	515 (23.4)
Visuo-spatial working memory	Posttest	1270 (24.4)	1199 (23.1)
	Follow up	337 (24.1)	333 (23.8)
Nonverbal abilities	Posttest	2713 (22.4)	2677 (22.1)
	Follow up	440 (23.2)	395 (20.8)
Verbal abilities	Posttest	1102 (29.0)	1138 (29.9)
	Follow up	112 (22.4)	136 (22.7)
Word decoding	Posttest	606 (37.8)	480 (30.0)
	Follow up	147 (29.4)	124 (24.8)
Reading comprehension	Posttest	759 (30.4)	650 (26)
	Follow up	84 (28)	84 (28)
Arithmetic	Posttest	1100 (37.9)	1088 (37.5)
	Follow up	393 (24.6)	377(23.6)

Figure s1a. Effect sizes at posttest for nonverbal ability, treated controls, for each study (displayed by ■) with confidence intervals, and mean effect size for treated controls (♦).

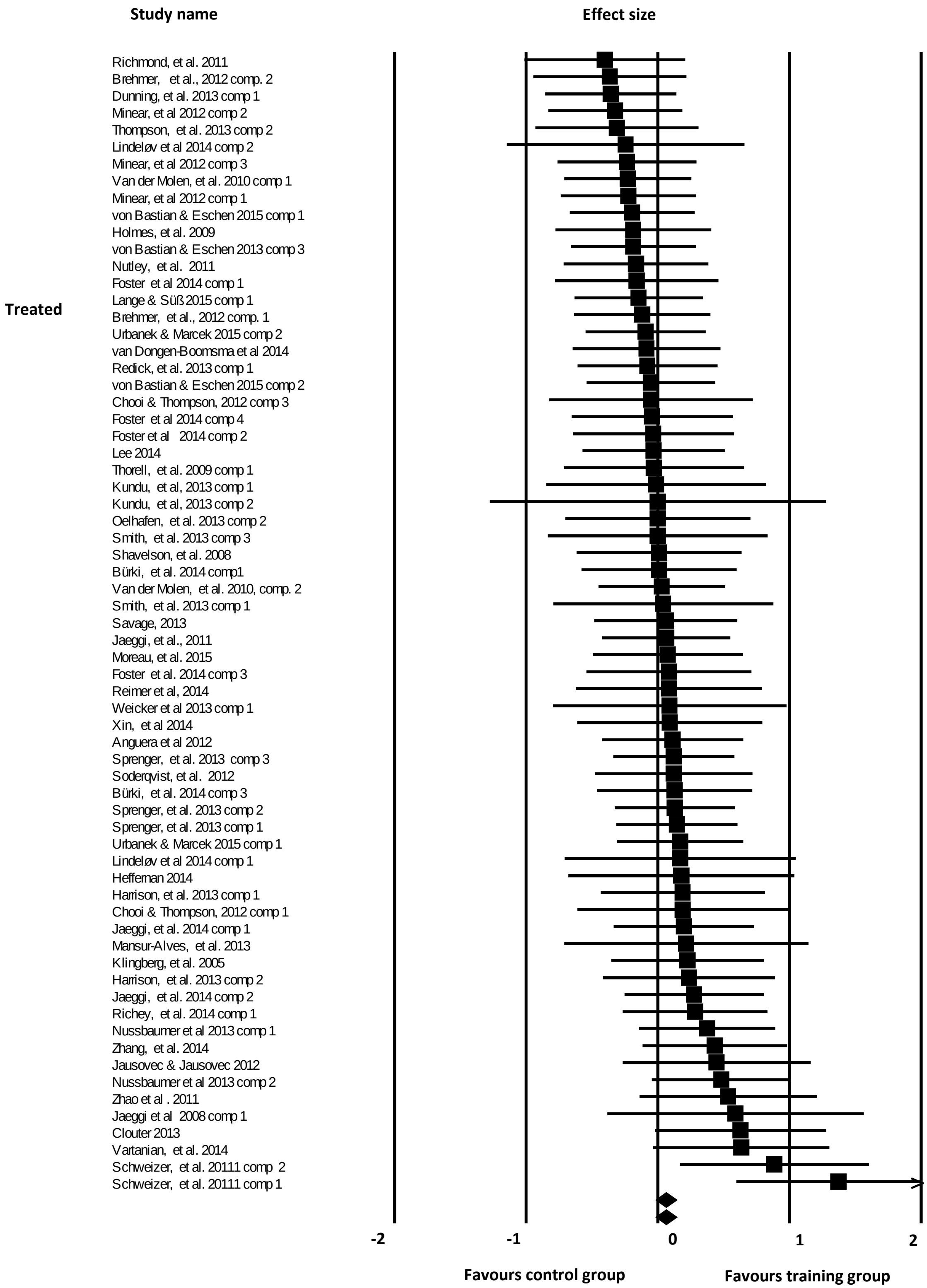

Figure s1b. Effect sizes at posttest for nonverbal ability, untreated controls, for each study (displayed by ♦) with confidence intervals, and mean effect size for treated controls (♦).

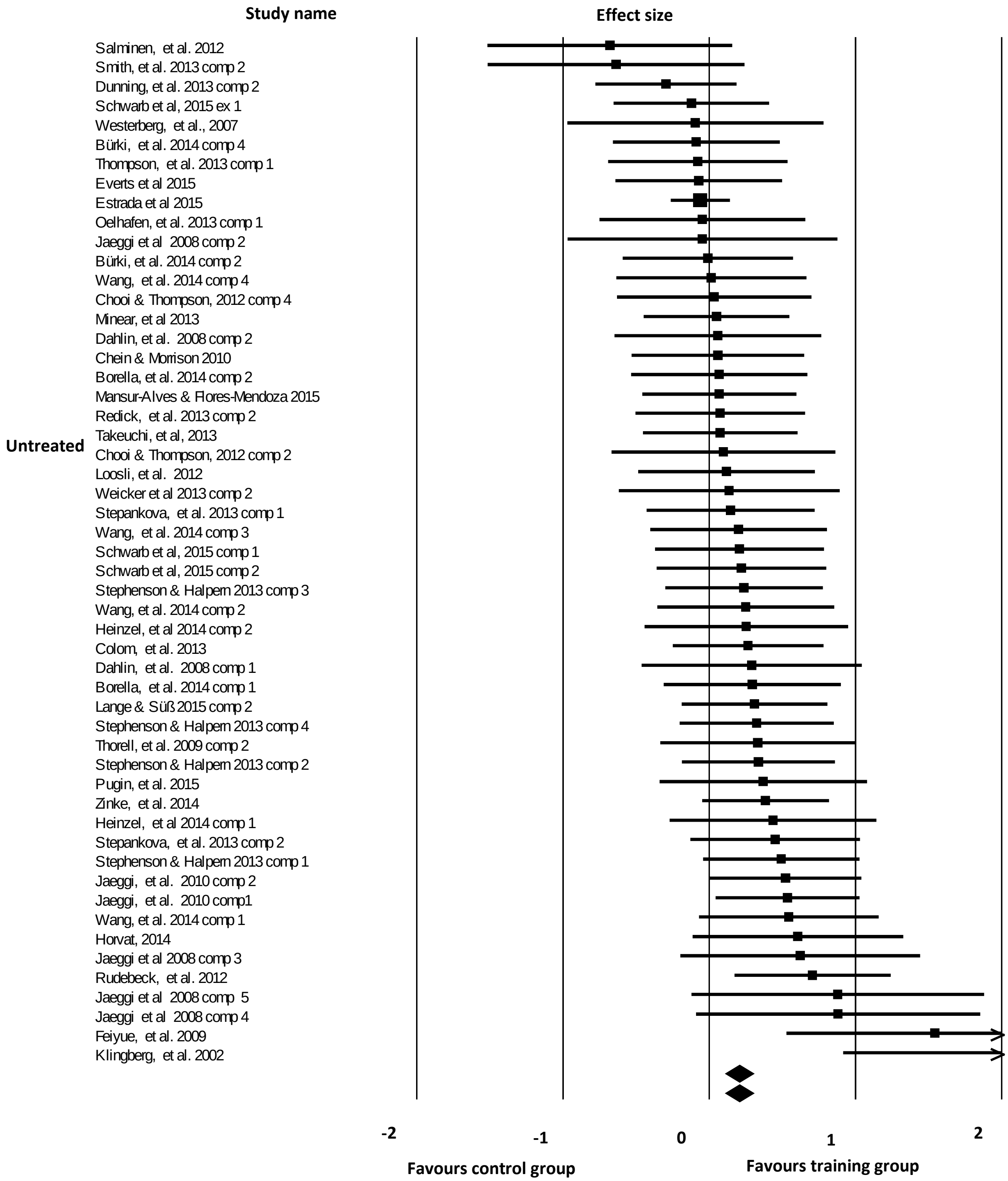

Figure s2. Effect sizes at posttest for verbal ability for each study (displayed by ■) with confidence intervals, and mean effect size for treated and untreated controls (♦).

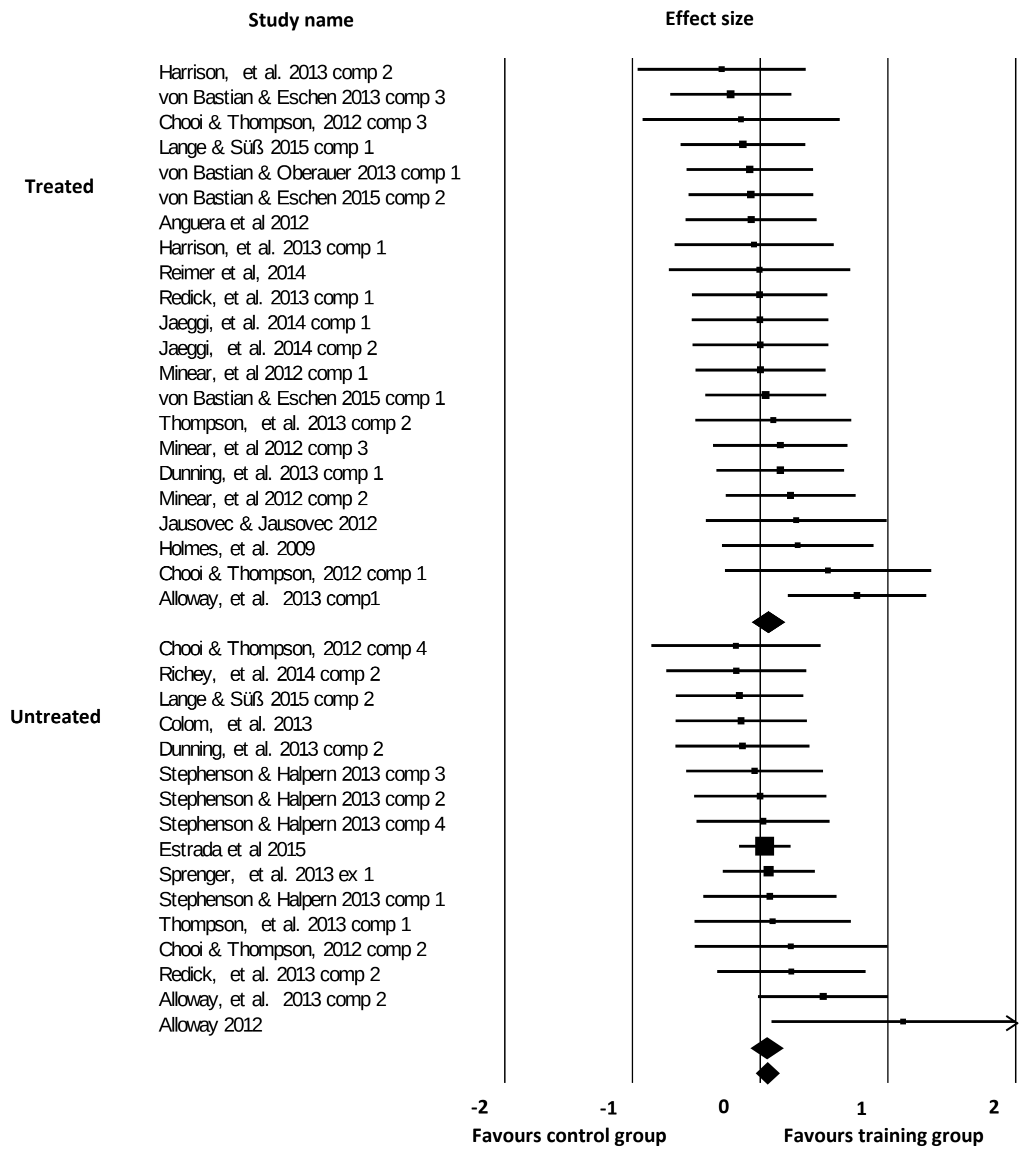

Figure s3. Effect sizes at posttest for decoding for each study (displayed by ■) with confidence intervals, and mean effect size for treated and untreated controls (♦).

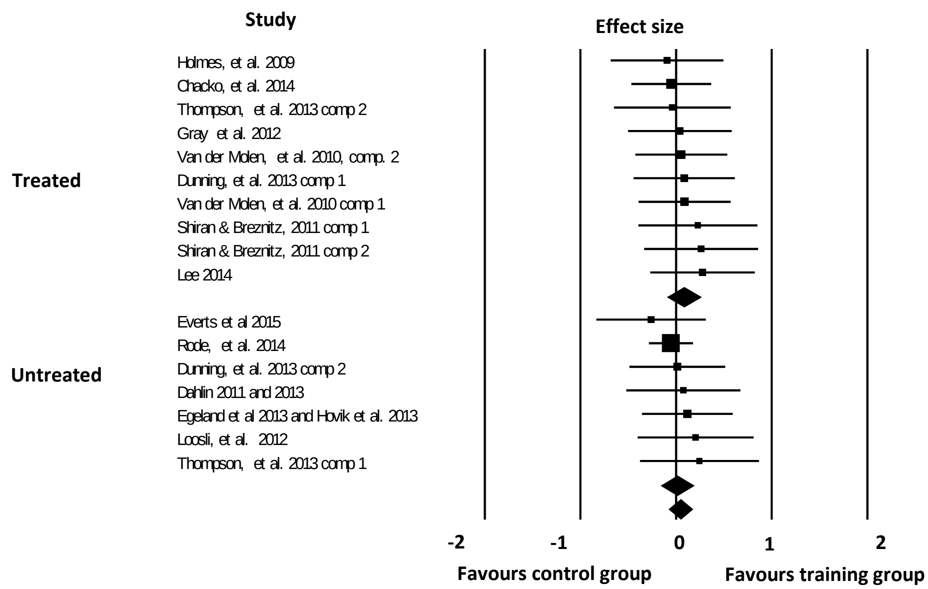

Figure s4. Effect sizes at posttest for reading comprehension for each study (displayed by ■) with confidence intervals, and mean effect size for treated and untreated controls (♦).

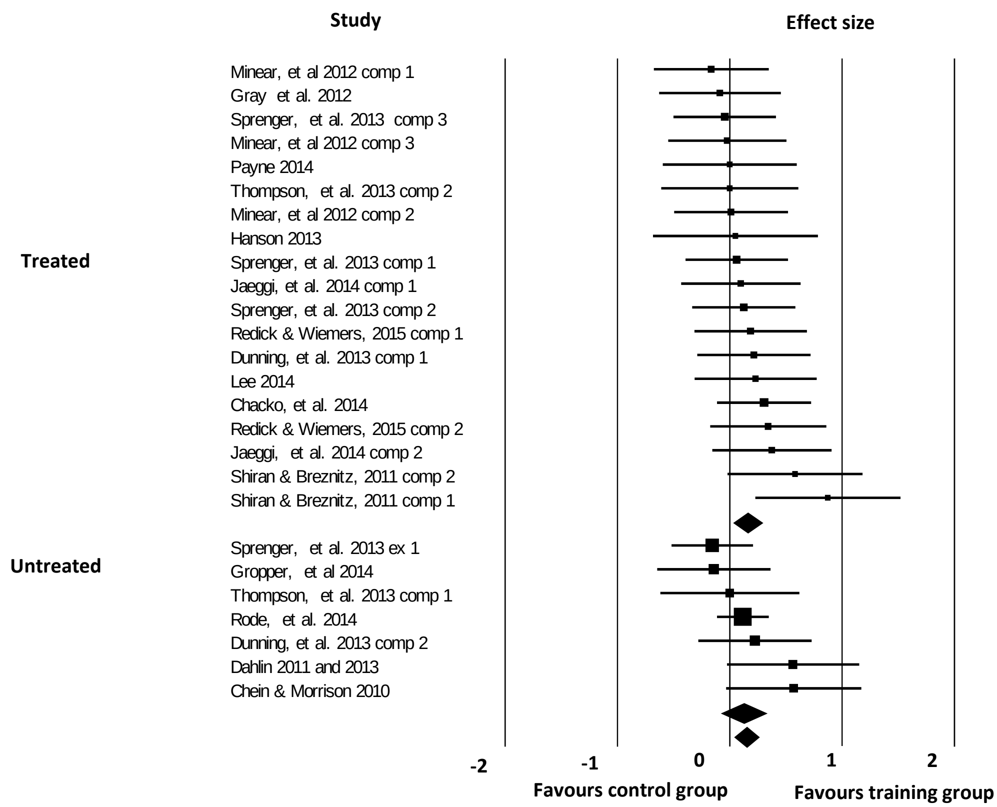

Figure s5. Effect sizes at posttest for arithmetic for each study (displayed by ■) with confidence intervals, and mean effect size for treated and untreated controls (♦).

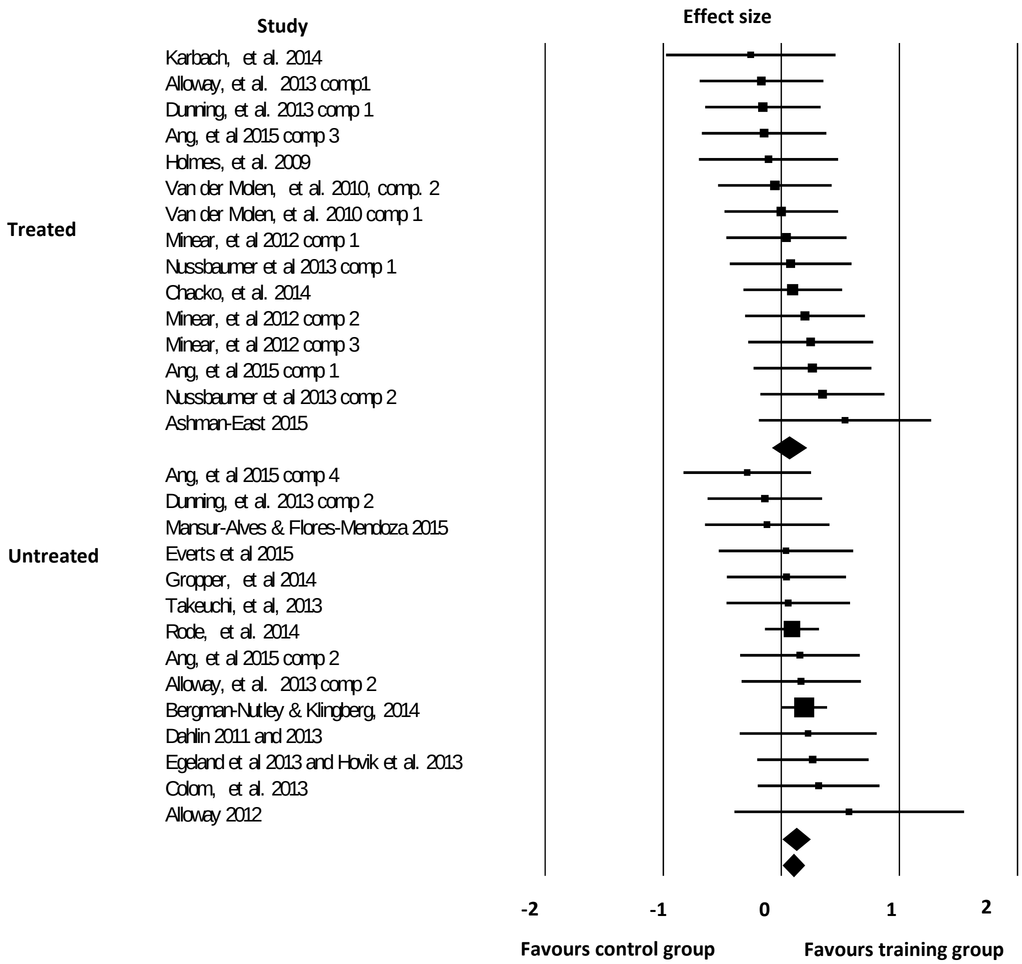

Figure s6. Effect sizes at posttest for verbal working memory for each study (displayed by ■) with confidence intervals, and mean effect size for treated and untreated controls (♦).

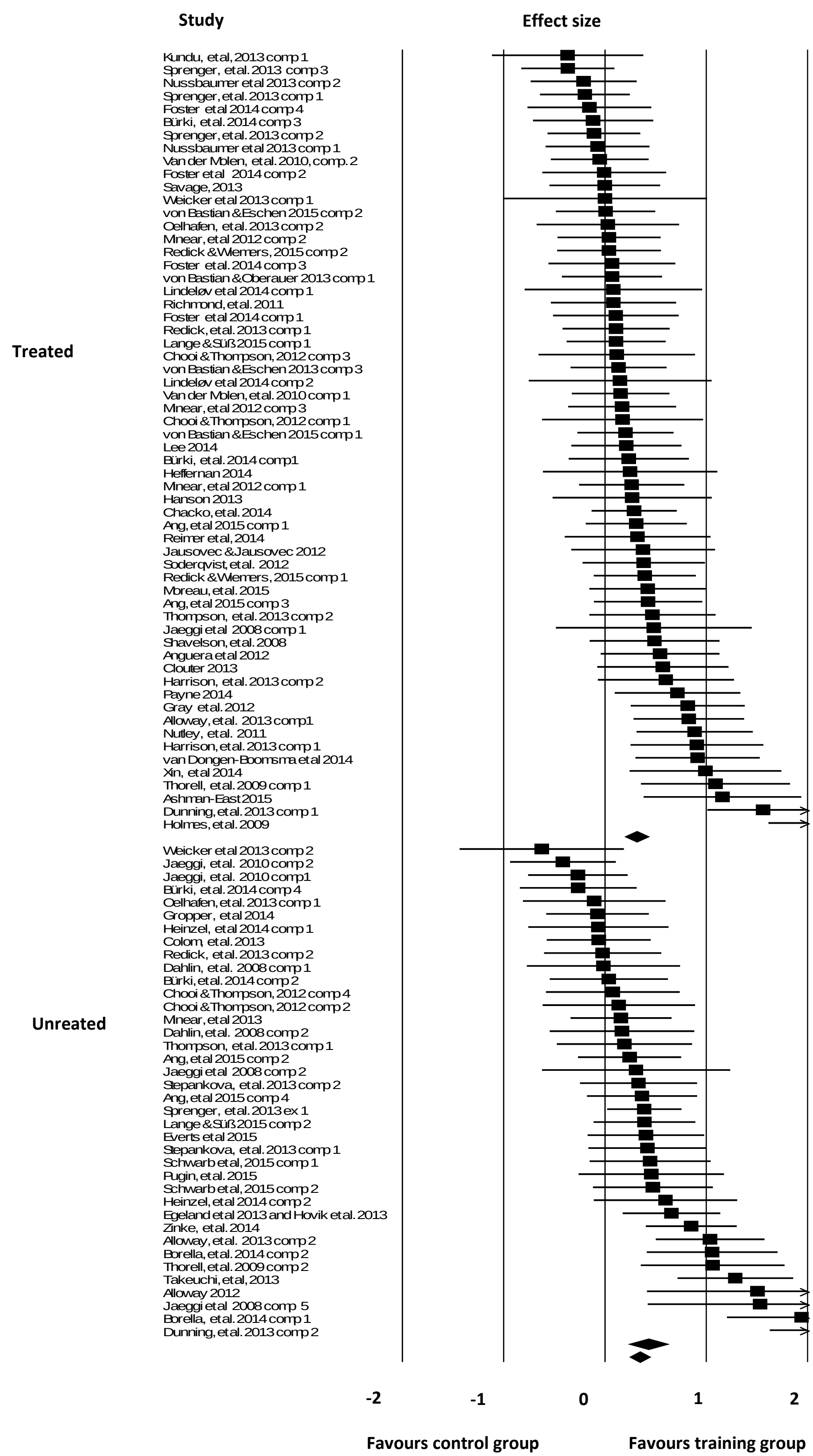

Figure s7. Effect sizes at posttest for visuo-spatial working memory for each study (displayed by ■) with confidence intervals, and mean effect size for treated and untreated controls (♦).

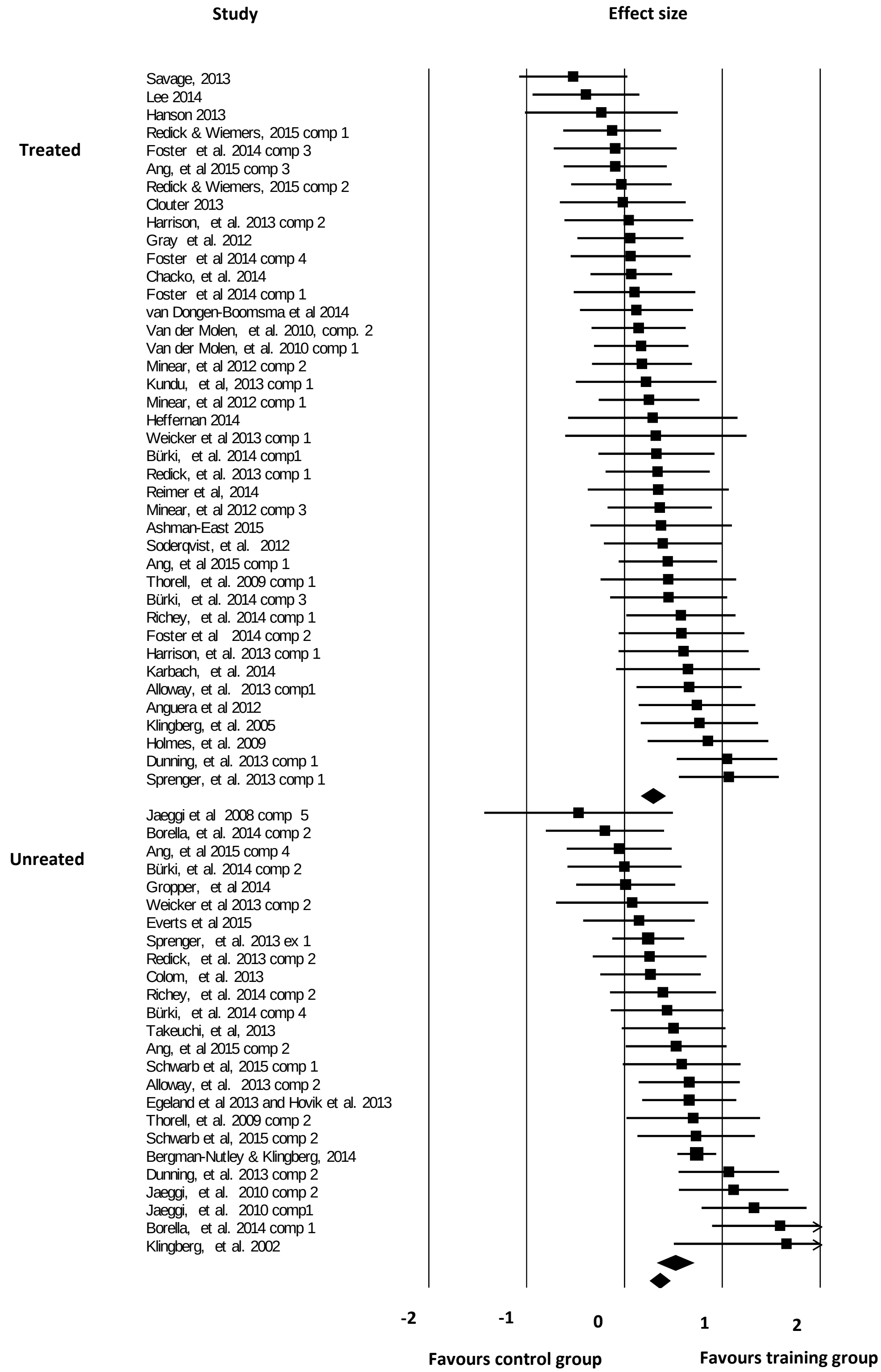

Figure s8. Effect sizes at posttest for criterion near transfer measures for each study (displayed by ■) with confidence intervals, and mean effect size for treated and untreated controls (♦).

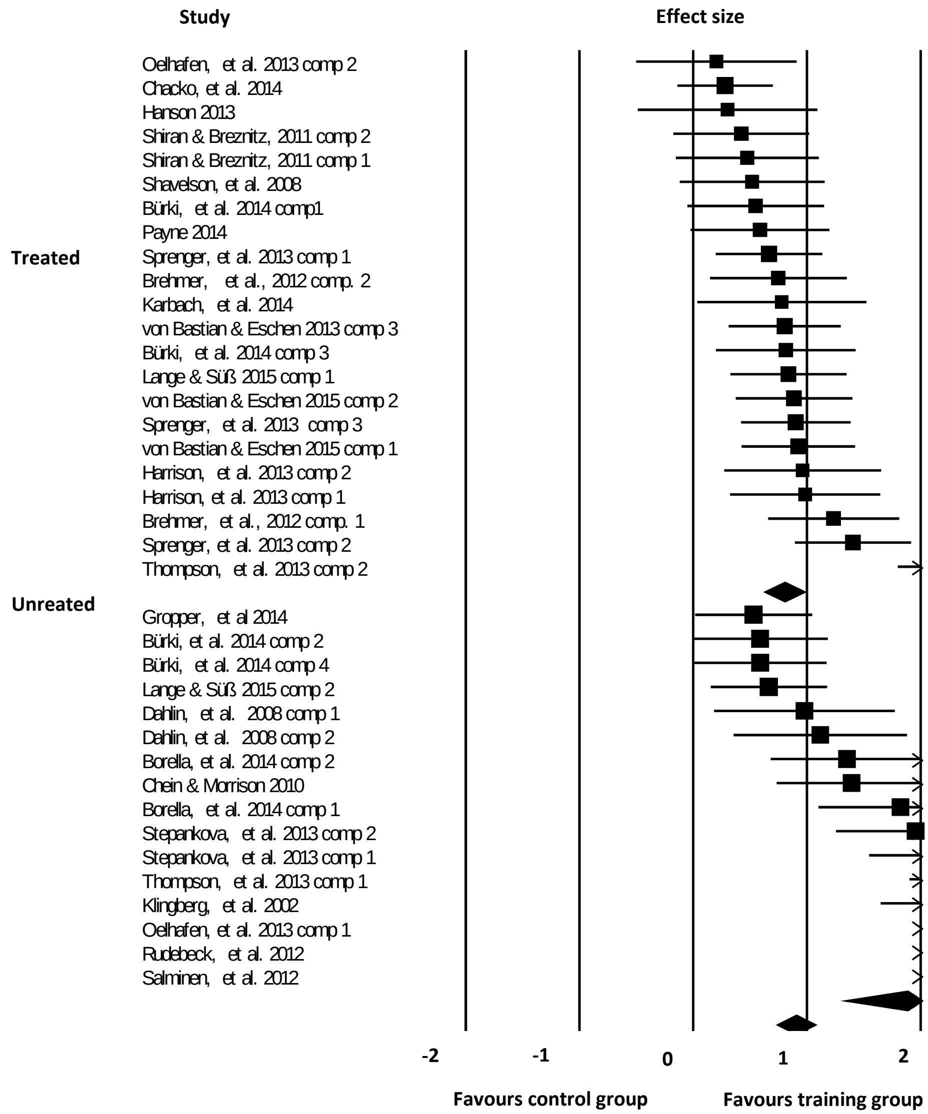

Table s4

Analysis of Moderators of Immediate Effects on Nonverbal Abilities

Moderator variable	Treated controls: Number of Effect sizes *(k)*	Effect size *(g)*	Heterog-eneity *(Tau²)*	Test of difference *(Q-test)*	Untreated controls: Number of Effect sizes *(k)*	Effect size *(g)*	Heterog-eneity *(Tau²)*	Test of difference *(Q-test)*
Age								
Children	14	-0.02	0		12	0.23	0.06	
Adults	47	0.10*	0		31	0.20**	0.02	
Older Adults	6	-0.13	0		10	0.22*	0	
				.13				.96
Training dose								
Large	29	-0.04	0		15	0.07	0.03	
Small	38	0.13**	0		38	0.23**	0.01	
				.02*				.14
Design								
Non-randomized	24	0.03	0		22	0.21*	0.02	
Randomized	43	0.07	0		31	0.19**	0.01	
				.64				.88
Learner status								
Learning disabled	11	-0.08	0		4	0.25	0.43**	
Unselected	56	0.08*	0		49	0.19**	0	
				.11				.88

Intervention program								
CogMed	9	-0.13	0		4	0.38	0.51	
N-back	30	0.15*	0		30	0.26**	0.02	
Complex span	13	-0.09	0		4	0.13	0	
Other	15	0.16*	0		15	0.11	0	
				.01*				.35

Table s5

Analysis of Moderators of Immediate Effects on Verbal Abilities

Moderator variable	Treated controls				Untreated controls			
	Number of Effect sizes *(k)*	Effect size *(g)*	Heterog-eneity *(Tau²)*	Test of difference *(Q-test)*	Number of Effect sizes *(k)*	Effect size *(g)*	Heterog-eneity *(Tau²)*	Test of difference *(Q-test)*
Age								
Children	-	-	-		-	-	-	
Adults	18	0.01	0		12	0.00	0	
Older Adults	-	-	-		-	-	-	
				-				-
Training dose								
Large	16	0.05	0.0		10	0.03	0.01	
Small	6	0.06	0.0		6	0.03	0	
				.97				.95
Design								
Non-randomized	12	0.12	0		7	-0.04	0	
Randomized	10	-0.02	0.01		9	0.05	0	
				.28				.53
Learner status								
Learning disabled	-	-	-		-	-	-	
Unselected	19	0.00	0		13	-0.01	0	

				-				-
Intervention program								
CogMed	-	-	-		-	-	-	
N-back	9	0.04	0		8	0.02	0	
Complex span	6	-0.03	0		-	-	-	
Other	5	0.13	0.10		6	0.09	.03	
				.71				.74

Table s6

Analysis of Moderators of Immediate Effects on Verbal Working Memory

Moderator variable	Treated controls				Untreated controls			
	Number of Effect sizes *(k)*	Effect size *(g)*	Heterog-eneity *(Tau²)*	Test of difference *(Q-test)*	Number of Effect sizes *(k)*	Effect size *(g)*	Heterog-eneity *(Tau²)*	Test of difference *(Q-test)*
Age								
Children	17	0.68**	0.21**		9	0.85**	0.30**	
Adults	37	0.12*	0		19	0.17	0.09	
Older Adults	6	0.28	0.07		10	0.49**	0.28**	
				.001**				.01**
Training dose								
Large	32	0.33**	0.16**		17	0.48**	0.25**	
Small	28	0.27**	0.02		21	0.37**	0.23**	
				.61				.57
Design								
Non-randomized	20	0.47**	0.21**		19	0.36*	0.32**	
Randomized	40	0.23**	0.05*		19	0.48**	0.15**	
				.08				.53
Learner status								
Learning disabled	17	0.58**	0.22**		8	0.76**	0.42**	
Unselected	43	0.18**	0.02		30	0.33**	0.17*	

				.01**				.11
Intervention program								
CogMed	11	0.91**	0.24*		5	0.84*	0.64**	
N-back	18	0.17**	0		19	0.12	0.03	
Complex span	16	0.19*	0		-	-	-	
Other	15	0.16	0.07**		12	0.48**	0.14**	
				.01**				.01*

Table s7

Analysis of Moderators of Immediate Effects on Visuo-spatial Working Memory

Moderator variable	Treated controls: Number of Effect sizes *(k)*	Effect size *(g)*	Heterog-eneity *(Tau²)*	Test of difference *(Q-test)*	Untreated controls: Number of Effect sizes *(k)*	Effect size *(d)*	Hetero-eneity *(Tau²)*	Test of difference *(Q-test)*
Age								
Children	17	0.32**	0.07*		9	0.62**	0.06*	
Adults	21	0.24**	0.05		12	0.44**	0.12**	
Older Adults	-	-	-		4	0.47	0.47	
				.73				.59
Training dose								
Large	21	0.35**	0.07		12	0.49**	0.07*	
Small	19	0.19*	0.02		13	0.52**	0.22* *	
				.15				.89
Design								
Non-randomized	13	0.42**	0.01		15	0.51**	0.10**	
Randomized	27	0.20**	0.06*		10	0.53**	0.16**	
				.051				.89
Learner status								
Learning disabled	16	0.30**	0.05		9	0.54**	0.09*	
Unselected	24	0.26**	0.05		16	0.49**	0.15*	

				.77				.77
Intervention program								
CogMed	10	0.34*	0.10*		7	0.60**	0.12**	
N-back	10	0.24*	0.03		9	0.52**	0.16**	
Complex span	11	0.18	0.02		-	-	-	
Other	9	0.37**	0.07		6	0.37**	0	
				.60				.67

Table s8

Analysis of Moderators of Immediate Effects on Criterion Near Transfer Measures

Moderator variable	Treated controls: Number of Effect sizes *(k)*	Effect size *(g)*	Heterog-eneity *(Tau²)*	Test of difference *(Q-test)*	Untreated controls: Number of Effect sizes *(k)*	Effect size *(d)*	Heterog-eneity *(Tau²)*	Test of difference *(Q-test)*
Age								
Children	4	0.41**	0		-	-	-	
Adults	14	0.90**	0.12**		8	2.27**	2.02**	
Older Adults	4	0.76**	0		7	1.37**	0.34**	
				.04*				.04*
Training dose								
Large	17	0.85**	0.11*		8	1.86**	1.22**	
Small	5	0.59**	0		8	1.91**	1.03**	
				.16				.93
Design								
Non-randomized	5	0.94**	0.47**		-	-	-	
Randomized	17	0.78**	0.02		13	2.03**	1.12**	
				.66				-
Learner status								
Learning disabled	4	0.37*	0		-	-	-	
Unselected	18	0.88**	0.07**		14	1.91**	1.06**	
				.01*				-

Intervention program								
CogMed	5	0.62**	0.08		-	-	-	
N-back	4	1.02*	0.74		8	2.52**	1.94**	
Complex span	6	0.84**	0		-	-	-	
Other	7	0.83**	0.03		-	-	-	
				.72				-

Table s9

Grey versus published studies

Construct	*Results*
Nonverbal ability	Results for published studies showed a mean *g* = 0.13 (95% CI [0.07, 0.19] *k* = 101) and for grey studies *g* = 0.04 (95% CI [-0.11, 0.18], *k* = 19). This difference was not significant ($p = 0.24$).
Reading comprehension	Results for published studies showed a mean *g* = .15 (95% CI [0.04, 0.27] *k* = 18) and for grey studies *g* = 0.08 (95% CI [-0.11, 0.28], *k* = 8). This difference was not significant ($p = 0.53$).
Verbal working memory	Results for published studies showed a mean *g* = 0.40 (95% CI [0.28, 0.52] *k* = 76) and for grey studies *g* = 0.19 (95% CI [0.05, 0.32], *k* = 22). This difference was significant ($p = 0.02$).
Visuo-spatial working memory	Results for published studies showed a mean *g* = 0.48 (95% CI [0.37, 0.60] *k* = 47) and for grey studies *g* = 0.06 (95% CI [-0.09, 0.20], *k* = 18). This difference was significant ($p < 0.01$).

Table s10

P-curve inclusion rules

Number	*Rule*
1.	All studies of working memory training published as journal articles showing a significant result after testing a hypothesis or a research question of effects on far transfer measures. Criteria regarding design and participants are the same as applied in the meta-analysis, with exception that non-computerized WM training also included here to increase sample size.
2.	In cases where paper report significant effects on more than one far transfer effect on the same comparison groups, the first and the last analysis was selected.
3.	A p-curve will be reported for studies that report only one p-value plus the *first* p-value that is reported in studies that report more than one p-value. Another p-curve will be calculated for studies that report only one p-value plus the *last* p-value that is reported in studies that report more than one p-value.

Table s11

P-curve disclosure table for the sample published articles

Original paper	**1. Quoted text indicating prediction of interest to researchers**	**2. Study design**	**3. Key statistical result**	**4. Quoted text from paper with conclusion based on statistical result**	**5. Results on far transfer measures (quoted statistical result)**	**6. Robustness of results**
Alloway 2012	*We postulated that any observed gains in vocabulary and academic attainment could be explained by the interactive training program rather than practice effects or test taking skills*	Randomised Untreated controls	Differences in means, nonparametric statistics	*The superior performance of the training group compared with the control group was confirmed in most of the cognitive measures: vocabulary, math and wm*	U = 8.5, p = 0.02 (vocabulary), in meta z = 2.12, p = 0.032	Math U = 12.5, p = 0.04. In meta,z score not sig.
Alloway et al 2013	*Specifically, we were interested in whether working memory training would result in transfer effects within an educational setting , measured by standardized tests of verbal ability and academic achievement*	Randomised Untreated controls and Randomised treated controls	ANOVA two way interaction	*For vocabulary there was not a significant difference in performance as a function of group or between times, but the interaction was significant , and the wm high group performed better than the wm low group post training.*	Excluded, since differences a baseline causes this result F(2,91) = 8.02, p = .001 vocabulary for treated control group	Follow up, F(2, 49) = 10.98, p = 0.001 vocabulary for treated control group
Ang et al 2015	*A question of continuing interest is whether academic performance can be improved by increasing WM or updating capacity. In this study, we designed and evaluated the efficacy of a computerized updating training programme.*	Non-randomised Untreated controls and Non-randomised treated controls	4 (Training condition) by 3 (assessment time) MANOVA	*Using the same MANOVA model, we examined the effectsof training on the mathematics tasks. Children improved acrosstesting sessions, but the magnitude of improvement was notaffected by training.*	Excluded, no significant far transfer effects	
Anguera et al 2012	*Type 2 tests included Raven's matrices (Raven et al., 1990), which is a standardized test of fluid intelligence, and the BOMAT and verbal analogies tests of intelligence (Hossiep et al., 1995). We have previously shown that working memory training transfers to*	Randomised treated controls	ANOVA	*A MANOVA with all the cognitive measures as dependent variables was significant (F(5,37) = 4.23, p < .05) showing more transfer overall for the NB group. Follow-up univariate ANOVAs revealed significant intervention*	F(5,37) = 4.23, p < .05	

	performance on this task (Jaeggi et al., 2008), and we included it here for the sake of replication.' Quote from Seidler et al 2010 technical report			*effects for the 3-back (F(1,41) = 4.68, p < .05), 4-back (F(1,41) = 4.70, p < .05), and operation span tasks (F(1,42) = 3.90, p < .05).*	
Bergman-Nutley & Klingberg 2014	*The inconsistent results of WM training on mathematics could be due to 19 a true lack of effect or that only certain aspects of mathematics are affected; 2) that effect occurs not directly after training but later as a combination of improved WM capacity in combination with instruction or 3) that the effect size is small, and that the existing studies include too few subjects to detect an effect.*	Non-randomised Untreated controls	General linear model and General linear model with repeated measures	*The training group improved significantly more than the control group on all three transfer tasks* (odd one out, following instructions and math)	Math F(1, 388) = 13.5, p = 0.0004
Borella et al 2014	*Concerning the transfer effects, we expected to find the same transfer effects, and maintenance effects, as were seen after administering the verbal WM training by Borella et al. (2010) to young–old.*	Randomised Untreated controls	Two way ANOVA with interaction	*Contrary to the results reported by Borella et al.'s (2010) verbal WM training study, no far transfer effects were apparent in our participants of either age group, with the exception of a processing speed measure showing that trained young–old completed tasks more quickly at the posttest stage, and this benefit was not maintained at the follow-up.*	Excluded, no significant improvement
Brehmer et al 2012	*Based on previous findings, we expected (a) younger and older adults to benefit from WM training, (b) near-transfer effects to non-trained WM tasks but also some far-transfer to tasks that share similar underlying pro- cesses (i.e., attention, reasoning), and (c) maintenance effects for younger as well as older adults across the 3-month time interval for the training gains as well*	Randomised treated controls	Mixed repeated measure ANOVAs were conducted with age (young and old) and intervention (adaptive training and low-level practice) as between-subject factors and time (baseline, post-training, and follow-	*Regarding far-transfer, similar performance improvements for the adaptive training as well as the active control groups were observed for tests of interference control (Stroop) and reasoning (RAVEN). These findings demonstrate general test-retest effects. More interestingly, both younger and older adults receiving adaptive*	F (2, 192)= 3.22, *p* = 0.045 cognitive functioning questionnaire

	as for potential transfer effects.		up) as within-subject factor for the eight cognitive tasks and the self-rating scale, respectively.	*training showed larger performance gains in a test measuring sustained attention (PASAT) and reported less memory complaints (CFQ) after the 5 weeks of intervention than the controls.*	
Bürki et al 2014	Note: Main aim with paper is to propose a model for analyzing individual learning curves in intervention research, not test hypotheses regarding far transfer. *To illustrate the proposed approach, a latent growth curve model analysis using data from a 10 day working memory training in younger and older adults is reported. .*	Non-randomised Untreated controls and Non-randomised treated controls	A repeated measures analysis of variance (ANOVA) was conducted including age-group (younger, older) as a between-subjects factor and training session (session 1–10) as a within-subjects factor.	*The fact that younger and older adults exhibited similar transfer effects is in line with some training studies (..) but contradicts others (..) which reported transfer effects in younger adults but not in older adults .No additional transfer effects, that is, transfer to other tasks, were observed. This result is in line with recent WM training studies (..) in which far-transfer effects were not reported, either in younger or in older adults.*	Excluded, no significant far transfer effects
Chacko et al 2014	*Moreover, given the relationships between working memory, inattention, and academic achievement, it was hypothesized that, compared to CWMT Placebo, CWMT Active would result in significant improvements in ADHD inattention symptoms, objective measures of attention, and academic achievement*	Randomised treated controls	Mixed effects regression was used for each outcome over time using SuperMix software	*Both treatment groups improved with treatment on measures of academic achievement, with no incremental benefit of CWMT Active on these outcomes. These findings are similar to those of Gray et al. (2012), who found no incremental benefit of CWMT Active compared to an intensive math intervention on academic achievement outcomes. This suggests that CWMT per se may not have specific effects on measures of academic achievement, at least in the short term.*	Excluded. No significant differences were found between treatment conditions on Word Reading, Sentence Completion, Math Computation or Spelling achievement scores at posttreatment.
Chein & Morrison	*We anticipated that, because CWM*	Randomised	T-test	*These WM training benefits*	t(38) = 1.80, p =

2010	*tasks place a strong demand on mechanisms linked to domain-general attention control (Engle & Kane, 2004), a training paradigm built around this task would result in both increases of WM span and more far-reaching benefits.*	untreated controls		*generalized to performance on the Stroop task and, in a novel finding, promoted significant increases in reading comprehension. The results are discussed in relation to the hypothesis that WM training affects domain-general attention control mechanisms and can thereby elicit far-reaching cognitive benefits.*	0.04 for reading comprehension (one-tailed)	
Chooi & Thompson 2012	*The present study predicted that there would be no improvements in verbal and perceptual tests, but there could be improvements in spatial ability and matrix reasoning tests.*	Non-randomised Untreated controls and Non-randomised treated controls	Paired t-test analyses	*Results from the current study did not suggest improvement in general intelligence after repeated training on a challenging working memory task. Our prediction that spatial and reasoning abilities could be improved after working memory training was not supported.*	Excluded. No significant improvement	
Colom et al 2013	*The main prediction is that if adaptive working memory training promotes skills relevant for the reliable temporary storage of relevant information, then fluid intelligence and working memory scores will be higher for the trained than for the control group at the posttest evaluation.*	Non-randomised Untreated controls	ANCOVA where the group was the independent variable, the construct/measure was the dependent variable, and the covariate was the score at the pretest for the corresponding variable. A *p* level of .05 (one-tailed) was considered for testing the results	*The main finding is that the large improvements in the challenging adaptive cognitive training program based on the* N-back *task (Fig. 2) do not evoke greater changes than those observed for a passive control group in fluid-abstract intelligence and crystallized intelligence, or in working memory capacity and attention control at the construct level.*	Excluded. No significant improvement RAPM]$F(1,53) = 2.340$; $p = .06$ (One-tailed ANCOVA)	
Dahlin 2011	*We hypothesized that working memory ability would increase through the training with a positive effect on children's reading comprehension skills (cf. Cain et al., 2004).*	Non-randomised Untreated controls	Multivariate analysis of variance with repeated measures	*The results show that working memory can be seen as a crucial factor in the reading development of literacy among children with special needs, and that interventions to improve*	Only the results of reading comprehension improved at T2, (estimated treatment effect	T3 reading comprehension: T 2.72, $p < .05$

				working memory may help children becoming more proficient in reading comprehension.	= 2.51, SE = 0.8, t = 3.27, p = .01, d = 0.88)
Dahlin 2013	*It was hypothesised that WM training at school for a period of five weeks would improve skills in WM ability, and subsequently improve results in mathematics.*	Non-randomised Untreated controls	Repeated measures model	*The results indicate that boys aged 9 to 12 with special needs may benefit, over time, from WM training, as shown in the enhanced results in mathematics following WM training.*	Treatment effect = 1.638, SE = 0.690, $F(1, 26) = 5.63$, $p < .05$
Dahlin et al 2008	*First, at both the group and the individual level, we examined whether young and older adults would improve their updating performance after updating training. Of main interest was whether older adults would show evidence of executive plasticity at all.*	Randomised Untreated controls	2 (Group: trained, control) x 2 (Session: pretest, Posttest 1) analyses of variance (ANOVAs) with repeated measures on the last factor were performed	*Transfer effects were in general limited and restricted to the young participants, who showed transfer to an untrained task that required updating (3-back).*	Excluded. No significant improvement For the two fluency tasks and reasoning, all interactions involving group and session were nonsignificant (*ps*> .05).
Dunning et al 2013	*Whilst improvements in WM tasks that closely resemble the trained activities are reported consistently (…), the evidence for transfer to tasks that share little overlap with the structure and content of trained activities while drawing on hypothesized common processes is mixed (..).*	Non-randomised treated controls and non-randomised untreated controls	To test group effects on training gains, general linear models were performed separately for the different T2 measures with scores at T2 entered as the dependent variable and scores at T1 and group as independent variables	*Adaptive WM training did not significantly improve children's performance on standardized reading and mathematics tests either immediately after training or one year later. Indeed, the only significant change in any group was an increase in basic reading scores for the no intervention group. It also had no effect on nonverbal reasoning, contrary to studies that have used N-back training paradigms (..), and others in which CWMT (…) has been used despite comparable statistical power. We therefore*	Excluded. No significant improvement

				have no evidence to support claims that WM training enhances nonverbal IQ.		
Egeland et al 2013	*Thus, in this analysis of far transfer effects we ask whether the increased WM performance transfers 1) to other NP functional domains, i.e. selective attention, sustained attention or learning capacity; (2) to academic skills such as mathematics and reading ability, and (3) whether parents and teachers rate the training children as less symptomatic with regard to a) working memory, b) attention in general, and c) ADHD symptoms.*	Randomised untreated controls	Treatment effects are analyzed applying Multivariate Analysis of Covariance (MANCOVA) with treatment condition as between group factor and PT1 and PT2 scores as within group factor. Pretest scores were entered as covariates.	*Reading and mathematics were improved. Text reading became faster and more correct. Decoding of single words became more correct, although not faster.*	F = 7.19, p < .001, df (1,59) for word decoding % correct	F = 2.34, p < .016, df (1,59) for word decoding % correct
Estrada et al 2015	*This brings to life the well-known practice effect and it must be taken into account in research aimed at the proper assessment of changes after the completion of cognitive training programs… Between the pre-test and the post-test sessions, some participants completed eighteen practice sessions based on memory span tasks, other participants completed eighteen practice sessions based on processing speed tasks, and a third group of participants did nothing between testing sessions.*	Randomized untreated controls	Nested SEM models fit separately for each group	*The good fit shown by thismodel implies that the three groups were equal before and after practice, meaning that differential practice was not associated with neither any differential effect in the latent variable weight on the test, nor the tests' means.*	Excluded, no significant far transfer effects	
Everts et al 2015	*This study aimed to determine whether two types of memory training approaches resulted in an improvement of trained functions and/or a generalization of the training effect to non-trained cognitive domains.*	Non-randomised untreated controls	Nonparametric tests of short-term and long-term gains; tests were computed one-sided and a significance level of p<0.05 was assumed.	*Children following a program of working memory training presented a significant improvement in trained functions (verbal working memory, visual short-term memory). Non-trained functions did not improve after the training.*	Excluded, no significant far transfer effects	
Feiyue et al 2009	*One of the issues which academic*	Non-	t-test of gain scores	*Through using Raven's*	No significance	

	people concentrates on is whether Gf of adults can be improved.	randomised untreated controls		*Standard Progressive Matrices as the evaluation method to get and analyze the experimental results, it was proved that training pattern can improve fluid intelligence of adults. This will promote a wide range of applications in the field of adult intellectual education.*	testing in paper; Correspondence t=4.785，p<0.000).
Gray et al 2012	*It was also anticipated that WM training would be associated with concomitant improvements in behavioral symptoms of ADHD in the classroom, with greater effects on inattention compared with hyperactivity/impulsivity. It was expected that WM training would be associated with subsequent improvements in those aspects of numeracy and literacy that are dependent upon WM (e.g., reading comprehension, math reasoning, and spelling) and that math training would be associated with improvements on math tasks.*	Randomised treated controls	Group differences were tested by comparing outcome (posttest) scores between the two groups using a between-group analysis of covariance (One-way analyses of covariance [ANCOVA]), with age and baseline score as covariates.	*In contrast with previous studies of WM training (..), we did not find robust evidence of improvements in behavioral symptoms of inattention or academic attainment.*	Excluded. No significant improvement
Gropper et al 2014	*also anticipated that WM training would be accompanied by improvements in academic areas that are dependent upon WM (e.g., reading comprehension and math reasoning), and improved self regulation in everyday life (albeit perhaps as later-onset outcomes discernible at follow-up only).*	Randomised untreated controls	ITT analyses used ANCOVA with baseline as a covariate and Group (experimental, control) as a between-subjects factor. The dependent variables were post-test scores on target indices.	*Computerized WM training is a feasible and possibly viable approach for enhancing WM in college students with ADHD or LD.*	F (1,59) = 4.39 cognitive failures questionnaire
Harrison et al 2013	*We also assessed transfer effects to tasks dissimilar to our training tasks but that were theorized to reflect WMC (which would indicate*	Randomised treated controls	ANCOVA with group as the between subjects variable and subjects'	*The results suggest that WMC and Gf are different hypothetical constructs and that an intervention that may*	Excluded, no significant improvement

	moderate transfer). Far transfer would be demonstrated if training on complex span tasks led to improvement on a battery of Gf tasks.		pretest performance as a covariate.	*improve WMC may have no effect on Gf.*	
Heinzel et al 2014	*Since processing speed and executive functions were expected to improve through our training approach, we expected to find a transfer effect to fluid intelligence. we expected younger adults to outperform older adults in training gains and transfer effects in the current study.*	Randomised untreated controls	2 (training vs. control group) × 2 (t1 vs. t2) ANOVAs	*Results suggest that working memory training may be a beneficial intervention for maintaining and improving cognitive functioning in old age. A significant group by time interaction was only found in younger adults, indicating improved performance in the Verbal Fluency test in the younger training group compared to the younger control group. Contrary to our hypothesis, no transfer to our speeded tasks of fluid intelligence (LPS Figural Relations Test and Raven's SPM) was found in the current study.*	$F(1, 28) = 5.55$, $MSE = 68.27$, $p = .026$, verbal fluency test
Holmes et al 2009	*The purpose of the present study was to answer these three questions by evaluating the extent to which the training program boosts performance of children with low WM on a standardized battery of untrained and well-validated WM tasks (…) and on measures of academic ability, both immediately following completion of training and 6 months later.*	Non-randomised treated controls	ANOVA group by time interactions	*This study provides the first demonstration that these commonplace deficits and associated learning difficulties can be ameliorated, and possibly even overcome, by intensive adaptive training over a relatively short period: just 6 weeks*	Excluded. No significant effects on far transfer measures. Mathematics 6 months after training $F(1, 17) = 9.50$, $MSE = 48.66$, $p < .01$., but no control group
Jaeggi et al 2008	*The aim of the training intervention was the investigation of the effects of training on the working memory task and its impact on Gf.*	Non-randomised untreated controls	Group X test-session interaction, test version as covariate	*The improvement in the groups that received the apparent benefit of training was substantially superior. The*	$F(1,67) = 5.27$; $p = 0.05$;

				finding that cognitive training can improve Gf is a landmark result because this form of intelligence has been claimed to be largely immutable. Instead of regarding Gf as an immutable trait, our data provide evidence that, with appropriate training, there is potential to improve Gf.	
Jaeggi et al 2011	*Nevertheless,it seems that Gf is malleable to a certain extent as indicated by the fact that there are accumulating data showing an increase in Gf-related processes after cognitive training (6). Referring back to the analogy in the physical domain, we can characterize WM as taking the place of the cardiovascular system; WM seems to underlie performance in a multitude of tasks, and training WM results in benefits to those tasks.*	Non-randomised treated controls	group × session (post vs. pre)	*However, despite the experimental group's clear training effect, we observed no significant group × test session interaction on transfer to the measures of Gf [group × session (post vs. pre): F(1, 59) < 1; P = not significant (ns); (follow-up vs. pre): F(1, 53) < 1; P = ns; with test version at pretest (A or B) as a covariate] (Table 1). Next, we compared transfer to Gf between these two training subgroups and the control group. Our results indicate that only those participants above the median in WM training improvement showed transfer to measures of Gf [group × session (post vs. pre);* F*(2, 58) = 3.23;* P *< 0.05 (Fig. 4*A*), with test version at pretest (A or B) as a covariate]. Planned contrasts revealed significant differences between the group with the large training gain and the other groups (P < 0.05; see Fig. 4*A *for effect sizes)*	$F(2, 58) = 3.23$; $p < 0.05$
Jaeggi et al 2014	*Since we had reason to believe that the processes underlying N-back*	Non-randomised	Univariate ANCOVAs for both	*This study incorporated several methodological advances*	Visuospatial Reasoning: F(2,

	performance are domain-free (..), we hypothesized that transfer to reasoning should not depend on the specific stimuli used in the training task. Finally, and most importantly, we used multiple fluid reasoning tasks that we combined into composite scores as transfer measures in order to investigate whether the effects that we had found previously were test specific, or whether the effects were more general on a construct level.	treated controls	composite gain scores (with Intervention Type as a between-subjects factor and test version as a covariate)	*over previous WM training studies, and nonetheless replicated transfer to measures of fluid intelligence (…). In particular, this study showed transfer to a composite score representing five visuospatial reasoning measures. Thus, transfer effects do not seem to be restricted to a specific task such as the BOMAT; rather, they seem to be more general, in that they emerged with respect to a visuospatial reasoning factor that did not consist of matrix reasoning tasks alone. Second, this transfer was observed despite the use of an active control group that trained on a knowledge-based task (which showed no improvements in visuospatial reasoning).*	74) = 3.51; p = .035 Single n-back vs controls [F(1, 50) = 7.20, p = .005 **one-tailed,** two tailed p = 0.01 Dual n-back vs controls Visuospatial Reasoning [F(1, 49) = 3.07; p = .04 **one-tailed**	
Jaeggi et al 2010	*Considering the rationale that transfer is more likely to happen for tasks that share considerable variance, we can conclude that training on both single and dual N-back tasks should yield transfer to matrix reasoning, but that transfer to working memory capacity should be less likely, especially in the case of single N-back training.*	Non-randomised untreated controls	Repeated-measures ANOVAs with session (pre vs post) as a within-subject factor, and intervention (dual nback, single N-backcontrol) as a between-subject factor separately for each matrix task (BOMAT and APM)	*But most interestingly, our results show transfer effects in both matrix reasoning tasks after training. This replicates our prior results (..), but it also extends our findings by showing that a) the transfer effect was present in more than just one Gf task, and b), that it was also obtained by training on a single N-back task.*	1. BOMAT: F(2,85) = 3.45; p = .05	2. APM: F(2,85) = 5.03; p = .01
Jausovec & Jausovec 2012	*The aim of the present study was to investigate whether training of WM functions (short-term storage and processing components like control of attention and executive functioning)*	Non-randomised treated controls	General linear model (GLM) for repeated measures test/retest x type of task (digit-span, RAPM, spatial	*The analysis of behavioral data revealed a significant increase of performance in respondents of the working memory group. This increase was most*	F(1,27) = 6.66; p < .05	

	can improve performance on tests of fluid intelligence		rotation, verbal analogy) x group (working memory, active control).	*pronounced for the RAPM, but also present on the other three test-batteries used. In conclusion, the results obtained, beside the mentioned limitations due to sample structure and size, lend further support to the hypothesis that working memory training can improve fluid intelligence which is also reflected in changed brain activity.*	
Karbach et al 2014	*To summarize, recent findings indicated that cognitive training may indeed support specific aspects of school-related abilities and academic performance in childhood. However, previous studies were mostly restricted to clinical subgroups (…) and it is unknown whether their findings generalize to healthy children. Therefore, the present study was designed to extend previous findings by testing the effects of adaptive training with a complex WM span task on academic abilities in the domains of math and reading in a sample of healthy elementary-school children.*	Randomised treated controls	ANOVAs with the factors Group (training, control) and Session (pretest, posttest).	*In the domain of academic abilities, our data showed short-term transfer to reading ability but not to math ability. The benefits for reading were substantial (d' = 1.08) and extend the findings on healthy children from Loosli et al. (2012) by showing that transfer of adaptive WM training is also significant when the adaptive WM training is compared to an active control condition.*	Reading: Session and Group, $F(1, 26) = 5.546$, $p < .05$,
Klingberg et al 2005	*A previous preliminary study indicated that training of WM tasks can enhance executive functioning including WM, response inhibition, and reasoning in children with ADHD (Klingberg et al., 2002b). A major shortcoming of that study was the low number of subjects (n = 7 in both the treatment and the comparison groups).The current study was therefore conducted at four*	Randomised treated controls	Hypotheses were tested by comparing outcome score at later times (T2 or T3) for the two groups using a general linear model, controlling for age, number of days of program use, and baseline score (T1). This analysis is	*The three other executive tasks (digit-span, Stroop task, and Raven's task) were secondary outcome measures, and the outcome of the statistical tests for these tasks should therefore be interpreted cautiously. However, group differences for Raven's task and the Stroop task were also found in the preliminary study of children*	Raven's task $n = 44$ $R2$ explained by total model = 0.77 beta = 2.1, $p = 0.01$ **One-tailed.** Not significant in meta-analysis

	clinical sites evaluating the effects of practice of WM tasks in a randomized, controlled, double-blind design.		equivalent to a between-group analysis of covariance with baseline as a covariate.	*with ADHD (..) as well as in a study of WM training in adults (…). Together, these results indicate that the effect of WM training also transfers to nontrained executive tasks other than WM tasks.*	
Klingberg et al 2002	*In the present study we investigated whether WM capacity could be improved by training. Furthermore, if impairment of WM is a core deficit in ADHD, this would imply that improvement of WM would decrease the symptoms in ADHD.*	Randomised untreated controls	Only p value, no sign test reported. *Significant improvement on Raven's Progressive Matrices was also evident (Table 1).*	*The improvement on the reasoning task is a clear evidence of that the training effect generalized to nonpracticed tasks, since the training did not include any problem solving or reasoning exercises at all. The improvement in reasoning ability is likely due to the fact that complex reasoning depends on WM, or more precisely, that the trained WM tasks and the reasoning task rely on the same cortical areas.*	p = .001 after t test reported in paper, no t-value reported. In meta: z = 3.72
Kundu et al 2013	*The aim of the present study was to investigate the neural bases of WM training effects, and of their transfer to untrained tasks. We trained an experimental group of subjects on an adaptive, visuospatial N-back task that has been shown to improve performance on other WM tasks, as well as on tests of fluid intelligence (…) and of reading comprehension*	Randomised treated controls	ANOVA group x session	*Both groups improved, in terms of accuracy and RT, on the DD and TD variants of the location VSTM task (Fig. 1d,e; Table 5). Notably, there was also an absence of WM training transfer to tests of complex WM span (Operation Span), fluid intelligence (RAPM), and control of response conflict (Stroop task)*	Excluded. No significant far transfer effects.
Lange & Süß 2015	*Thus, transfer to short-term memory, speed, and reasoning was only expected after near transfer was found.*	Randomised treated controls and Randomised untreated controls	2 (pretest, posttest) × 3 (training group, active control group, passive control group) mixed ANOVAs	*Although there were significant training effects, no transfer effects were found.*	Excluded, no significant far transfer effects
Loosli et al 2012	*To conclude, evidence for improved*	Non-	Multivariate analysis	*Concerning the transfer*	The MANOVA

	reading after WM training is very scarce, and there are no studies available investigating whether WM training improves reading processes in typically developing children. Therefore, the goal of the current study is to determine whether WM training in this group will lead to transfer effects to important school-related domains; in our case, reading performance. In addition to reading performance, we also included a transfer task, which is highly correlated with measures of scholastic achievement, namely Gf as measured with a matrix reasoning task.	randomised untreated controls	of variance (MANOVA) with group experimental, control) as the between factor and the differences between post- and pretest scores (from this point on termed gain scores) as dependent variables. Pillai's V as an *F*-statistic	*measures, we found an overall larger performance increase in the experimental group as indicated by the MANOVA. The MANOVA was driven by the gain in reading performance, that is, in reading of text and words but not pseudowords. Finally, in contrast to many previous WM training studies, which often looked at transfer on other laboratory tasks, we showed that it is possible to improve an ability that is very important in everyday life and is related to scholastic achievement in school-aged children.*	with all outcome measures (*Gf* , reading of pseudowords, words, and text) as dependent variables was significant, $F(4, 35) = 3.80$, $p < .05$
Mansur-Alves et al 2013	*The present research intends to verify the effectiveness of a cognitive training (CT) to foster intelligence of school Brazilian children from different intellectual levels.*	Randomised treated controls	Wilk's lambda	*no statically [sic] significant difference was found between both groups at posttest in none of the measures used*	Excluded, no significant far transfer effects
Mansur-Alves & Flores-Mendoza 2015	*Recent investigations applying working memory training have indicated that it is possible to train intelligence. This work aimed to verify the effectiveness of a cognitive training program aimed at increasing children's intelligence.*	Randomised untreated controls	MANOVA with group (experimental, control) as the between factor and the standardized change (using the formula: post-test – pre-test/SD pre-test) as dependent variables was used	*The statistical analysis indicated no signif cant differences between EG and CG after training for cognitive measurements. These results demonstrate partial support of the selective literature that indicates the difficulty of achieving significant intellectual changes through specific intervention programs.*	Excluded, no significant far transfer effects
Moreau et al 2015	*Based on prior research in working memory training using complex span tasks, we predicted working memory — but not spatial ability — gains after training working memory.*	Randomised treated controls	Separate 3 (Condition) × 2 (Session) mixed factorial ANOVAs with repeated measures on the latter	*Simples effects conductedwith dependent t-tests showed improvements for all groups, yet of different magnitudes. The DS group showed the largest improvements, followed by the*	Excluded, no significant far transfer effects for WM training versus AE

			variable were conducted for each task	*WM group and the AE group.*	
Nussbaumer et al 2013	*In sum, the main goal of the current study is to test a) whether a WM training yields near transfer, an enhancement of performance in untrained WM tasks, and b) to systematically test whether such a potential WM enhancement can provoke far transfer in the domain of intelligence and mathematical problem solving and whether such an enhancement is depending on the amount of WM load during training.*	Treated randomised	ANOVA (between subject factor group: low, medium and high load and within-subject factor time: pre-, post-, and follow up- testing)	*No differential transfer occurred in any of the mathematical problem solving tasks or in the intelligence tests. Positive transfer occurred between two tasks focusing on inhibitory processes.*	Excluded, but $F(2,79) = 3.31$ $p < 0.05$ (inhibition)
Nutley et al 2011	*The main aims of this study were therefore to investigate: (1) if Gf is improved through computerized training on non-verbal reasoning (NVR) tasks; and (2) if training WM or NVR would result in any transfer to measures of the non-trained construct, Gf and WM, respectively.*	Treated randomised	The expected value of the latent variable from T2 (given the tests scores) was used as a dependent variable in an ANCOVA with group as fixed factor, age in months and the expected value of the latent variable at T1 as covariates. In the event of a significant or marginally significant ($p < .10$) group effect, planned comparisons were performed	*In summary, we found that Gf can be improved through5 weeks of NVR training in 4-year-olds. This type of training might be useful for children with poor intelligence. Early detection and intervention of children who would benefit from NVR and or WM training could possibly prevent falling behind at school and allow learning opportunities that may otherwise be lost due to impaired cognitive capacities*	Excluded, no improvement in the WM group. ($F(3, 101) = 4.64$, $p = .005$ Planned comparisons revealed that the NVR training group ($p = .02$) and the Combined training group (wm + NVR) ($p = .05$) had improved significantly more than the placebo training group

Oelhafen et al 2013	*Our primary objective was to test whether training of shared cognitive processes in the training and transfer tasks would lead to improved performance in the ANT and corresponding electrophysiological changes.*	Randomised Treated controls and untreated controls	Behavioral data were analyzed with analyses of variance (ANOVA), and for pairwise comparisons, we conducted Tukey HSD (within subject) and Games-Howell (between subject) corrected tests	*However, the main effect of group and the relevant session by group interaction did not reach significance, Fs<1. Also, combining the two training groups and comparing them to the passive control group did not reveal a main effect of group or a group by session interaction for RST and BOMAT (all Fs<1). Thus, the RST and BOMAT scores were higher in the posttest, but neither the lure training nor the non-lure training group showed a higher pre–post gain compared to the control group.*	Excluded. No significant transfer effects
Pugin et al 2015	*The aim of our study was to investigate working memory training and its effects on working memory tasks and fluid intelligence in male subjects between 10 and 16 years. In fact, this age range may be particular susceptible to interventions because many cognitive functions are still developing. Furthermore, working memory performance has been shown to be linked to attentional control[10] and processing speed[11]. Thus, putative transfer effects on fluid intelligence may not be limited to* fluid intelligence, but may also include other cognitive functions.	Non-randomised Untreated controls	A mixed ANOVA test between 'group' and 'test session'	*A mixed ANOVA test revealed a significant difference between 'group' and 'test session' in auditory* N-back *(ANB) performance. No other test (letter-number sequencing task, number-span task, matrix reasoning task, Stroop task, and Flanker task) showed a significant change.*	Excluded. No significant transfer effects (Open access journal)
Redick et al 2013	*Numerous recent studies seem to provide evidence for the general intellectual benefits of working memory training. In reviews of the training literature, Shipstead, Redick, and Engle (2010, 2012) argued that the field should treat recent*	Non-randomised, treated and untreated groups	Factorial ANOVAs with Group 3 as the between subjects factor and Session 3 as the within-subjects factor. Significant Group x Session	*Despite improvements on both the dual* N-back *and visual search tasks with practice, and despite a high level of statistical power, there was no positive transfer to any of the cognitive ability tests.*	Excluded. No significant transfer effects (Out of 17 ANOVAs, there were no significant

	results with a critical eye.		interactions were decomposed with simple effects analyses focusing on the effects of Group and Session independently.		Group x Session interactions)
Richey et al 2014	*This study investigated whether working memory improvements, if replicated, would increase analogical reasoning ability. We assessed participants' performance on verbal and visual analogy tasks after a complex working memory training program incorporating verbal and spatial tasks [3], [4].*	Non-randomised untreated controls	Paired-samples t-test for each group	*Participants' improvements on the working memory training tasks transferred to other short-term and working memory tasks, supporting the possibility of broad effects of working memory training. However, we found no effects on analogical reasoning.*	Excluded. No significant transfer effects (open access journal)
Richmond et al 2011	*We predicted that: (a) older adults would show improved WM span after training, and (b) older adults would show far transfer to assessments of everyday functioning. In addition, we predicted a replication of the finding in Chein and Morrison (2010) that this particular WM training paradigm does not produce far-transfer to a common measure of general intelligence, Raven's Progressive Matrices.*	Treated randomised	Chi-square test of Likert scale concerning cognitive functioning, ANOVA on continuous data	*Compared to the trivia control subjects, a significantly greater number of participants in the training group selfreported an increase in attention when queried about general cognitive improvements they thought may have been affected by training*	χ^2 (1, *n*= 9) = 2.78, *p* = .05 (One-tailed**)** Cognitive functioning self report
Rode et al 2014	*At this point, it is however not clear to what degree such a core working memory training program can be embedded within a regular school context and to what degree it produces benefits on academically relevant abilities that exceed those of regular class participation. Therefore, the main goal of the current project was to adapt an existing training program [19] to and*	Untreated, Non-randomised	t-tests	*However, only the AWMA, the CBM Math, and the Teacher Rating showed significant condition differences.*	t (281) = 2.20 (math)

	test ifs effectiveness within a relatively large sample of 3rd grade students within a classroom context.				
Rudebeck et al 2012	*We predicted, therefore, that our spatial WM training task would at least lead to significant improvements in Gf and potentially, in recognition memory performance as well, as captured by some or all of the different performance measures used.*	Untreated randomised	To investigate any changes in performance after training, gain scores (post- minus pre-training score) were calculated for all tasks	*Overall, the trainers made a significantly greater improvement on this test [Bomat] in comparison to controls*	t(53) = 3.14, *p* = 0.003
Salminen et al 2012	*In summary, the present study set out to investigate, whether training effects from the dual* N-back*transfers to (1) a WM updating task, (2) dual-tasks with different demands on WM updating, (3) task switching, and (4) an AB task. Additionally, transfer to reasoning abilities was tested*.	Untreated randomised	2 (Group: training vs. control) × 2 (Session: pre-test vs. post-test) mixed-design ANOVA	*In any case, we provided no evidence for WM transfer effects to the performance in the RAPM after training.*	Excluded. No significant transfer effects (open access journal)
Schwarb et al 2015	*Researchers have promoted the enticing possibility that simple behavioral training can expand the limits of working memory which indeed may also lead to improvements on other cognitive processes as well*	Randomised untreated controls	ANOVA group by time interaction	E1: *In this experiment, n-back training did not improve Gf.* E2: *As in Experiment 1, in this experiment, Gf did not improve following WM training.*	Excluded, no significant far transfer effects
Schweizer et al 2011	*Our first hypothesis then was that training on the dual* N-back*task (irrespective of the valence of the content), relative to control task training, would lead to transferable gains in short-term memory/WM capacity (measured by digit span) and in Gf (measured by Raven's Progressive Matrices) over and above any gains in digit span.*	Treated randomised	ANOVA group by time interaction	*Our data provide some support for this by showing significant pre- to post-training improvements in Gf, ..the current results which further support the malleability of Gf to training have a potentially wide range of (encouraging) implications for educational, neuropsychological and psychopathology treatment settings, if they prove to be robust.*	Gf: group by time interaction was significant, *F*(1, 40) = 7.47, *p* = 0.01, There was a trend toward a significant group difference in G*f* (RPM scores) at pre-training, *p*≤0.10. After removing

					subjects that caused baseline differences *F*(1, 30) = 3.66, *P* = 0.032,
Shiran & Breznitz 2011	*The aim of the current study was to examine the effect of the CogniFit Personal Coach computerized training program on the recall range and speed of processing in working memory of dyslexic readers, and whether it affects reading ability.*	Treated non-randomised	ANOVA group x training	*The dyslexics' reading scores before memory training were significantly lower than those of the skilled readers for all reading measures except oral reading comprehension (Table 5). Following memory training, there was a significant increase in all measures except orthographic accuracy test. It can be concluded that our findings support the notion of plasticity in the neural system underlying working memory and point to a relationship between larger working memory capacity and enhancement of reading skills.*	Excluded. No significant transfer measures. Words per minute: F(1,28) = 1.18 Pseudowords per minute: F (1, 28) = 1.56 Not significant (paper emphasize only main effects from training not training x group interaction)
Smith et al 2013	*We expected that (a) there would be performance improvements for participants actively using training software, either commercial or custom-build, in comparison to any control groups and (b) that there would be increased improvement for participants using the custom-built training software, e.g. as in (Jaeggi et al., 2008).*	Randomised treated and untreated controls	A repeated measures analysis of variance6 with the between-subjects factor training intervention (Control, Gaming, COTS, DIY) and the within-subjects factor time point (pretest [week 0], posttest [week 3] and delayed posttest [week 4]) was conducted. The dependent variable was the RPM score at each time point.	*In the RPM tests both cognitive training systems (COTS/DIY groups) failed to produce significant improvements in comparison to the Control group or the Gaming group. This suggests caution in the over generalization on the effectiveness of brain training systems with results from other demographic groups, for example school children*	Excluded. No significant wm training effects on transfer measures. F(6,70) = 2.831, p = 0.016 on Raven, but the posthoc tests showed that the only significant change in the RPM score was between the posttest and

					delayed posttest in the Gaming group ((p = 0.017).)
Soderqvist et al 2012	*Second, we aimed to evaluate if successful training in children with intellectual disability leads to improved performance on non-trained tasks.*	Treated, non-randomised	To test the effect of training we performed univariate general linear using each of the outcome measures as a dependent variable and including T1 performance on the same measure, age, gender, group, and a group*gender interaction as independent variables.	*Training did not lead to significant improvements on reasoning ability tasks (Block Design and Raven's colored matrices) although a trend association was observed on improvements on Block Design for males*	Excluded. Effects of training were associated with improvements on Block Design in males with a trend effect [$F_{(1,17)}$ = 13.48, *p* = 0.062], No significant effects of training progress were observed for improvements on word span forwards, Raven's colored matrices or for Auditory Attention (all *p*-values >0.1).
Sprenger et al 2013	*The present paper addresses some of the shortcomings in prior studies. First, rather than focusing on a single training task, we evaluated the impact of training on a battery of training tasks. Our goal was to test the hypothesis that broad training yields broad transfer.*	Treated and untreated groups, randomised	ANCOVA testing for post-test differences between conditions controlling for pre-test performance; Bayes factor analysis	*Although participants showed improvement on the trained task and on tasks that either shared task characteristics or stimuli, we found no evidence that training led to general improvements in working memory. Using Bayes Factor analysis, we show that the data generally support the hypothesis that working memory training was ineffective at improving*	Using Bayes factor no support for transfer effects. Ordinary t-tests showed significant transfer effects on one of 6 far transfer variables in experiment 1

				general cognitive ability.	with untreated controls, deciphering languages, t (n = 55 C, n = 58 T) = 2.59, P = < 0.05. No significant transfer effects in experiment 2 with active controls.
Stepankova et al 2014	*The foremost goal of the current study was to examine the efficacy of an adaptive computer-based WM intervention in healthy, community-dwelling older adults. In addition to investigating transfer to WM and visuospatial skills, we were especially interested to see whether training frequency and training gain predicted the extent of transfer in both constructs, which would thereby extend previous findings demonstrating either a dose–response effects in Gf or a relationship between training gain and transfer on Gf.*	Untreated randomised	Univariate ANCOVAs using the posttest composite as dependent variable, the pretest composite as a covariate, and group (CG, Ex10, Ex20) as a between-subject factor. Helmert contrasts in order to compare transfer performance on the group level (i.e., CG vs. Ex10 *and* Ex20; Ex10 vs. Ex20).	*the present results add to the evidence for the malleability of visuospatial skills (Uttal et al., 2013) and, more specifically, to the few studies reporting transfer on visuospatial skills following WM training in older adults. To conclude, our data demonstrate generalizing effects to composite scores reflecting WM and visuospatial skills in young-old healthy adults after a verbal* N-back*intervention. Our work adds to the accumulating evidence for transfer effects in old adults by means of an easily accessible noncommercial computer-based program that can be used independently at home.*	Training on the N-back task resulted in improved visuospatial skills, $t(61) = 3.29$, $p = .001$ (one-tailed ANCOVA, $r = .39$) as compared with the CG.
Stephenson & Halpern 2013	*A number of theorists (e.g., ..) have viewed Gf as being a biologically predetermined ability. The results of Jaeggi et al., 2008 and Jaeggi et al., 2010 studies, however, have significant implications for the way philosophers, psychologists, and educators think about intelligence*	Randomised passive controls	A repeated measures analysis	*The primary goal of our study was to test the hypothesis that scores on tests of Gf would improve only for participants who had a visuospatial component in training to improve WMC. Overall, we found this hypothesis to be*	Raven Dual n-back vs passive controls, t(131) = 3.46, visual N-backvs passive controls t(131) = 2.80

	because the take-home message is that the ability to solve novel problems can be improved with a short training program. Their studies also have implications for the psychometric properties and uses of the APM and other tests of Gf. Therefore, a careful analysis is needed to substantiate and determine WMC mechanisms that are being improved and leads to improvement in Gf. The current study sought to determine what mechanisms in WMC might be improved through cognitive training, whether there are sex differences in the improvements, and determine the generalizability across other measures of Gf and cognitive tests.			*supported, but with a surprising finding that a visuospatial STM training program was also beneficial.*	
Takeuchi et al 2013	*We hypothesized that WMT would increase resting-FC within DMN, increase anticorrelations between DMN and EAS, increase resting-CBF in PFC, and increase rGMV in EAS. The hypotheses relating to resting-FC and resting-CBF are based on the abovementioned previous studies that showed that conditions with reduced WMC are generally characterized by a decrease in resting-FC within DMN, a decrease in anticorrelations between DMN and EAS, and a decrease in resting-CBF in PFC. The hypothesis relating to rGMV is based on our previous proposition described above.*	Untreated, non-randomised	Because the superiority of training was our primary interest, in our behavioral analysis, test-retest changes in the WMT group were compared to those in the control group using one-tailed one-way analyses of covariance (ANCOVAs) with the difference between pre- and post-test measures as dependent variables and pretest scores as independent variables (p < .05).	*WMT led to improved performance on RAPM but not on BOMAT. On the other hand, although several studies have showed the effects of WMT on Raven matrix tests (Takeuchi et al., 2010b), the effects of WMT on non-verbal reasoning fluid intelligence tasks have recently been contested and there may be a number of reasons for the lack of significant effects on BOMAT or other non-verbal reasoning tasks (Redick et al., in press; Takeuchi et al., 2010b).*	Compared with the control group, the WMT group showed significantly greater pre- to post-test increases in performance on (RAPM; p = .019, one-tailed ANCOVA)

			The use of one-tailed test is consistent with our previous study as well as those of other laboratories (Klingberg et al., 2005, 2002)		
Thompson et al 2013	*Recently, however, researchers have reported gains in fluid intelligence after multiple sessions of adaptive working memory training in adults. The current study attempted to replicate and expand those results by administering a broad assessment of cognitive abilities and personality traits to young adults*	Treated and untreated non-randomised	T-test, Bonferroni corrections	*The major finding was a failure to observe any gains in measured fluid intelligence after working memory training.*	Excluded, Participants did not generally show improvements on the tasks measuring near or far transfer (Published open access)
Thorell et al 2009	*We therefore hypothesized that both training programs would have effects on the trained construct, as well as show transfer effects to the other (i.e. WM would have effects on inhibition and vice versa). Furthermore, performance of both WM and inhibitory tasks requires continuous attention, and we therefore hypothesized that we would find transfer effects to laboratory measures of attention for both types of training.*	Non-randomised Treated and untreated	In another set of similar ANCOVAs (see Table 1), the two training groups were compared with the combined control group	*significant overall effect was, however, found for omission errors on the auditory CPT, as well as a marginally significant effect on omission errors on the go/no-go task. Planned comparisons revealed that the WM group, but not the inhibition group, had improved significantly more over time compared to the control group.*	Excluded But $F = 3.30$ *df* s(1, 24) for go no go inhibition
Urbánek & Marček 2015	*Previous research has shown mixed results for the ability of working memory training to improve fluid intelligence. The aims of this study were first to replicate these improvements…*	Randomised Treated controls	Repeated-measures ANOVA	*These tests revealed neither main effects…nor their interactions…for any of the intelligence measures.*	Excluded, no significant far transfer effects
Van der Molen et al 2010	*The focus of the current study is on adolescents with mild to borderline intellectual disabilities. This group is known to have substantial WM*	Treated, randomised	General linear model analysis (GLM) was performed controlling for baseline scores	*Scholastic abilities compound score increased for both group A and group B compared with the control group. In both cases,*	Beta = 0.14, p = 0.03 for training group A versus controls,

	problems (…), generally performs poorly on academic achievement domains (….) and requires more educational support than do typically developing adolescents (…). Given the relationship between WM performance and scholastic abilities, it is of substantial interest to study the feasibility and effectiveness of a WM training in adolescents with M-BID		(pre-testing).	*it was an increase in score on the Arithmetic test and not on the Reading test, which was responsible for the change... the results of the current study are encouraging in that apparently WM, a central and important cognitive aspect, can be trained effectively with a fanning out effect on scholastic and other everyday tasks in a cognitive weak and therefore vulnerable group of people; children with mild to borderline ID*	Beta = 0.16, P = 0.02 for training group B versus controls on scholastic abilities at follow up (arithmetic test)
van Dongen-Boomsma et al 2014	*WM shows a rapid development throughout preschool and early school-age (Carlson, 2005). Training children at this young age, before larger demands from school exist, could be beneficial by increasing WM capacity and thereby preventing development of cognitive and/or behavioural problems (Rueda, Posner, & Rothbart, 2005; Thorell, Lindqvist, Bergman, Bohlin, & Klingberg, 2009). Therefore, investigating the efficacy of WMT in younger children in ADHD is worthwhile.*	Treated, randomised	When all assumptions were valid, analysis of the Covariance (ANCOVA) was applied to optimize control for the variance at baseline. For each parameter, the endpoint measurement was the dependent variable, the baseline measurement a covariate, and groups the independent variable	*There was a significant difference between groups ($p < .001$) on the Start Index. Nevertheless, the WMG significantly improved on the task, as illustrated by a significant difference between the Start Index and the Max Index ($t(25) = 15.59$, $p < .001$).*	Excluded. No significant transfer effects. Both the active and the placebo condition improved on many outcome measures over time. However, no additional effect in favour of the active condition was found on any of the primary or other secondary outcome measurements
Vartanian et al 2013	*Compared to participants enrolled in an active control condition, we*	Randomised Treated	mixed-model ANOVA	*Second, compared to participants enrolled in the*	RAPM interaction:

	predicted that participants enrolled in the experimental condition would exhibit improvement in fluid intelligence	controls		*active control condition, we predicted that participants enrolled in the experimental condition would exhibit improvement in fluid intelligence. This prediction was confirmed as unlike participants in the active control condition who exhibited no change in fluid intelligence, participants in the experimental condition exhibited an 8% improvement in fluid intelligence at the end of training compared to baseline.*	$F(1, 32) = 5.90$, $p = .021$
von Bastian & Eschen 2015	*In the present study, we therefore tested the hypothesis that adaptive WM training is superior to other training procedures because task difficulty is continuously adapted to individual performance instead of being varied performance-independently, thus differentiating between adaptivity and variability of task difficulty. Hence, adaptive training was compared to another WM training procedure in which task difficulty varied randomly. In addition, a third WM training procedure was included in which participants themselves could modify training task difficulty. The purpose of this training procedure was to explore whether change in training task difficulty across the training period in the adaptive training condition approximately matches what the average individual would choose as the optimal modification of task difficulty across training. Finally,*	Treated, randomized	To evaluate gain from pre- to post-assessment, we computed standardized gain scores (i.e., difference between posttest and pretest score divided by the pretest standard deviation) for each individual and each task (cf. von Bastian & Oberauer, 2013). We then ran linear mixed effects (LME) models to estimate these gain scores on the level of generalization range (i.e., training, intermediate transfer and far transfer effects) rather than on the	*In summary, the results showed that adaptive WM training led to larger gains in the trained tasks than active control training. However, there was no consistent evidence for transfer to structurally dissimilar WM tasks or to reasoning tasks.*	Excluded. No significant results concerning far transfer.

	to evaluate whether we could replicate our earlier findings showing benefits after adaptive WM training on untrained, structurally dissimilar WM and reasoning tasks (von Bastian & Oberauer, 2013), we added an adaptive active control group solving trivia quizzes with low WM demand.		level of single tasks (for a more detailed discussion of the advantages of using LME models over analyses of variance		
von Bastian & Oberauer 2013	*With the present study, we wanted to answer the following questions: (1) can WMC (with its two aspects, stor- age and processin g, and relational integrati on) and supervision be improved by extensive training, (2) do training effects transfer to non-trained tasks measuring the same construct, and (3) does transfer to related cogni- tive abilities – such as inhibition and reasoning – occur?*	Treated, randomised	We therefore chose linear mixed-effect (LME) modeling, which is less. To assess whether the experimental groups differed in how much they gained in performance from pretest to posttest in different constructs, we modeled standardized gain scores (i.e., pre–post differences in performance, expressed in stan- dard deviation units) of each individual in each task as a function of experime- ntal group, construct measure d by the task, and their interaction.	*Storage-Processing training had an effect on working memory and reasoning, and Supervision training improved task shifting and reasoning.*	Storage processing vs active controls: Reasoning b = 0.21, *p* = 0.021
Wang et al 2014	*Based on the body of research on spacing, memory, and skill acquisition, we predicted that training schedule would have a sub- stantial impact on working memory training gain and transfer.*	Untreated, randomised	A Paired-sample *t*-test was performed for each of the four training groups together with the control group on	*We can draw two main conclusions from this study. First, training schedule has a significant impact on transfer of training. Second, the transfer effect of the 20-day group*	The SPM gain in the 20 Days group was significantly larger than the control group,

	Specifically, we predicted that the group(s) with the most spacing of training would improve most on the training task and further- more show the most transfer. In addition to this primary goal, we wished to replicate the results of other studies that have trained memory updating and found transfer to fluid intelligence in chil- dren (Jaeggi et al., 2011; Zhao et al., 2011). The total number of studies in which updating is trained in typically developing children is rather small, and thus this study provides an addi- tional data point with respect to the potential effects of updating training more generally.		SPM pre-test and post-test to evaluate the training transfer.	*replicated results of a recent study that used the identical training and transfer tasks (Zhao et al., 2011). It is also consistent with studies that have found improvements in typically developing children following working memory training*	$t_{(38)} = 1.832$, $p = 0.038$ (one-tail test)
Westerberg et al 2007	*By using the same method in this pilot study, we evaluated the effects of WM training on cognition in a group of patients with stroke*	Untreated, randomised	a general linear model and controlling for baseline scores. This analysis is equivalent of an ANCOVA analysis with baseline score as one of the covariates.	Furthermore, improvements were also significant on the nontrained tests for WM and attention (PASAT and RUFF 2&7),	Passat P = 0.001 Ruff P = 0.005 (no other information in the paper about f values)
Xin et al 2014	*We specifically focused on transfer to the following: (i) a running WM task using a different type of item as used during training (hereafter termed nearest transfer); (ii) a digit-span task forward and backward (near transfer); and (iii) performance on a measure of fluid intelligence (far transfer).*	Treated, non-randomised	ANCOVA	*The absence of a training effect on fluid intelligence in our study (far-transfer) is in accordance with previous studies with older adults adopting WM updating tasks during training and measures of reasoning ability during transfer tests (..).*	Excluded. No significant main or interaction effects for far transfer measuers maximum $F(1, 27) = 3.03$, $p = .09$,
Zhao et al 2011	*Thus, the present study used the running memory task to train WMU ability, enabling us to investigate whether training WMU specifically could directly promote the*	Non-randomised, Treated controls	Independent-samples *t*-tests of gain scores	*The analysis of the additive mean values revealed that fluid intelligence scores in the training group were significantly higher than those*	$t(31) = 2.271$, $P < 0.05$

	improvement of fluid intelligence.			*in the control group*	
Zinke et al 2012	*Considering the previous literature two possible predictions can be derived. On the one hand, it may be that especially in old-old adulthood only those participants who have maintained a certain cognitive status profit from a WM intervention that requires a considerable amount of attentional resources [18, 19, 20] . On the other hand, resting on the disuse hypothesis [21, 22] , it may be that participants who start with a low cognitive status (a possible result of a decline in using cognitive resources) are able to reactivate some of their potential with the help of training.*	Untreated, non-randomised	T tests	*Importantly, independent t tests revealed no significant differences between groups in percent gains in either Stroop interference scores or Raven*	Excluded. No significant far transfer effects. Raven scores (training group: M = 5.2%, SD = 11.5, control group: M = 11.5%, SD = 16.9; t(34) = –1.32, $p > 0.01$
Zinke et al 2014	*Taken together, the current study explored the limits and potential of WM plasticity in a sample of older adults ranging from young-old to old-old age. For that purpose, an experimental approach was used to compare a training group with a control group on measures of training and transfer performance. With an individual difference approach, possible moderating factors of training-related plasticity were investigated for training and transfer gains.*	Untreated, randomised	A two-factorial ANOVA was used with group (training vs.control group) as the between-subjects factor and time of measurement as the within-subject factor.	*For the fluid intelligence task (Raven SPM), there was a significant effect for the crucial interaction between the time of measurement and the group,*	$F(1, 78) = 5.0$, $p = .03$,

Note. The same inclusion criteria as used in the meta-analysis were applied to the *p*-curve analysis.