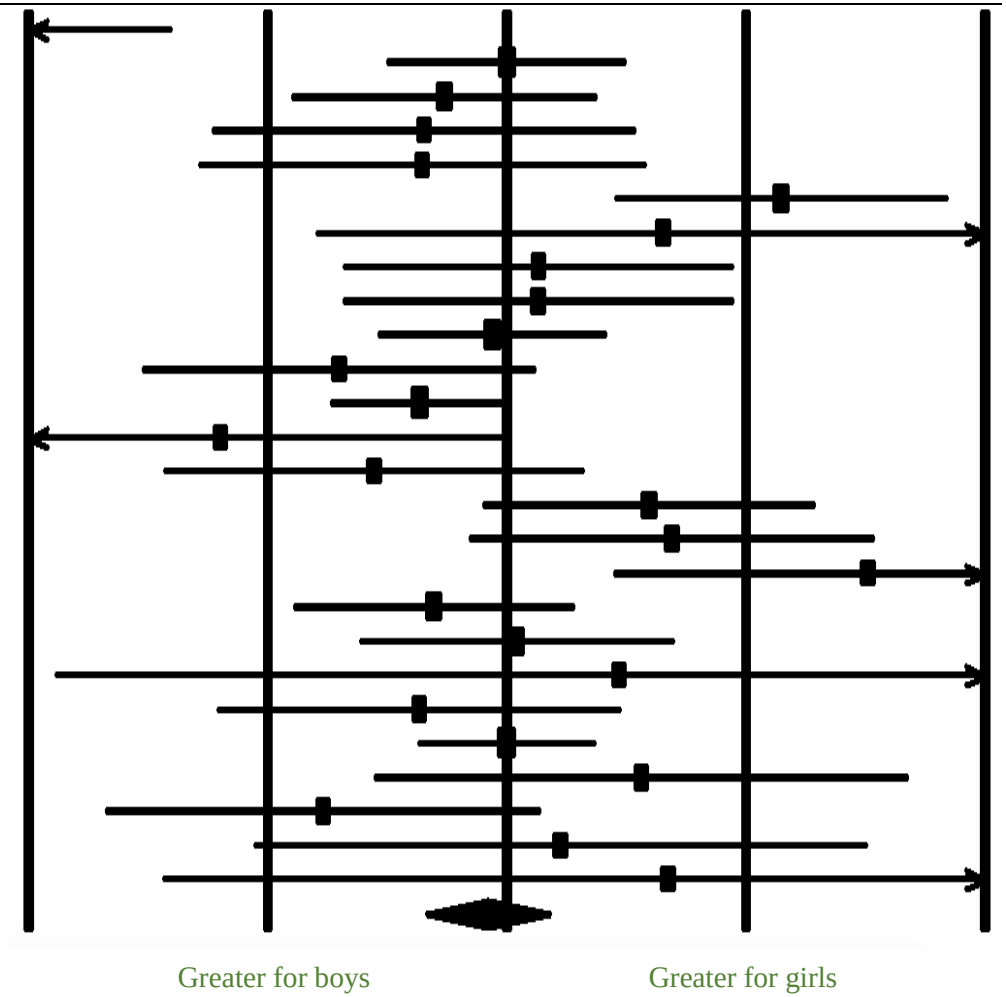
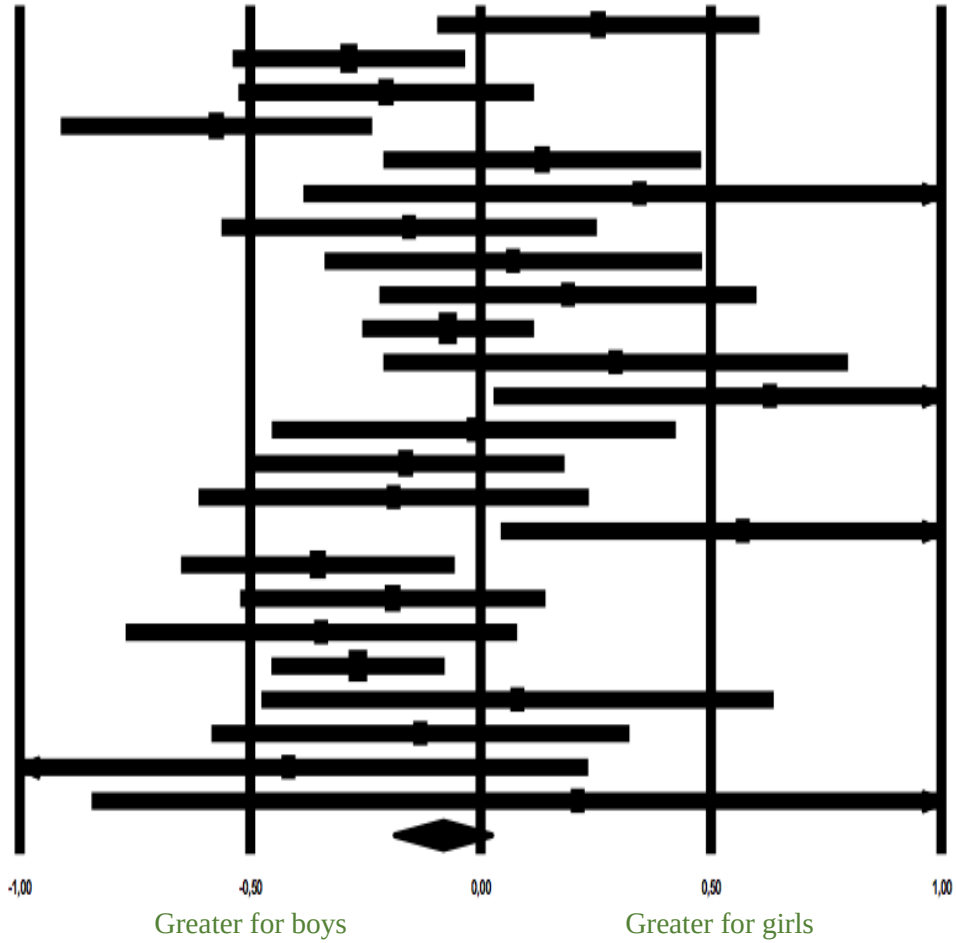
Study	Hedges'g	[95% CI]
DuPaul et al. (2006)	-1.064	[-1.427, -0.701]
Elkins et al. (2011)	0,000	[-0.248, 0.248]
Gökçe et al. (2017)	-0.130	[-0.447. 0.187]
Graetz et al. (2005), (2006)	-0.173	[-0.615, 0.268]
Hartung et al. (2002), Lahey et al. (2007)	-0.177	[-0.643, 0.290]
Heckel et al. (2009)	0.574	[0.227, 0.920]
Bilgiç et al. (2016)	0.327	[-0.397, 1.052]
Isaksson et al. (2016)	0.066	[-0.340, 0.472]
Lambek et al. (2010)	0.066	[-0.341, 0.473]
Mayfield et al. (2016)	-0.030	[-0.268, 0.207]
Montiel-N. et al. (2007)	-0.350	[-0.760, 0.059]
Muller et al. (2011a, 2011b)	-0.183	[-0.366, 0.001]
Nolan et al. (1999)	-0.598	[-1.194, -0.003]
Skogli et al. (2013)	-0.278	[-0.715, 0.160]
Riddle et al. (2013)	0.298	[-0.048, 0.643]
Rosch et al. (2015) & Seymour et al. (2016)	0.345	[-0.077, 0.768]
Rucklidge et al. (2001), (2002)	0.755	[0.225, 1.286]
Chan et al. (2014)	-0.152	[-0.444, 0.140]
Sonuga-Barke et al. (2007)	0.021	[-0.307, 0.348]
Sonuga-Barke et al. (2008)	0.234	[-0.943, 1.412]
Wang et al. (2015)	-0.183	[-0.604, 0.238]
Wehmeier et al. (2012)	0.000	[-0.184, 0.184]
Rucklidge et al. (2006)	0.281	[-0.275, 0.837]
Oie et al. (2016)	-0.384	[-0.836, 0.069]
Chen et al. (2008)	0.113	[-0.528, 0.753]
Coccini et al. (2009)	0.337	[-0.718, 1,392]
Summary	**-0.040**	[**-0.169, 0.090**]



Supplementary Figure 1 Inattention symptoms: Differences between boys and girls with ADHD

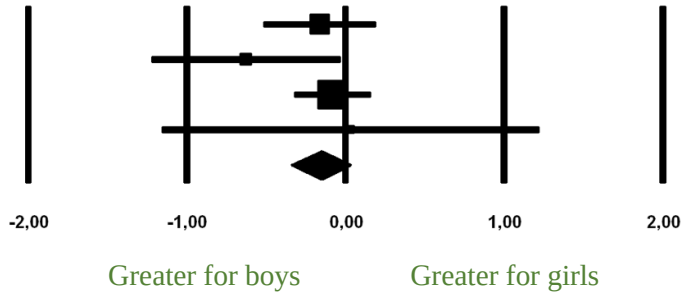
Note. CI= confidence interval. A positive effect size indicated more severe symptoms in girls, whereas a negative effect size indicated more severe symptoms in boys.

Study	Hedges'g	[95% CI]
DuPaul et al. (2006)	0.255	[-0.091, 0.602]
Elkins et al. (2011)	-0.286	[-0.535, -0.036]
Gökçe et al. (2017)	-0.205	[-0.522, 0.113]
Hartung et al. (2002), Lahey et al. (2007)	-0.573	[-0.908, -0.238]
Heckel et al. (2009)	0.134	[-0.208, 0.476]
Bilgiç et al. (2016)	0.345	[-0.382, 1.072]
Isaksson et al. (2016)	-0.155	[-0.559, 0.249]
Lambek et al. (2010)	0.071	[-0.336, 0.478]
Montiel-N. et al. (2007)	0.190	[-0.216, 0.596]
Muller et al. (2011a, 2011b)	-0.071	[-0.254, 0.113]
Borchuk et al. (2017)	0.293	[-0.208, 0.794]
Nolan et al. (1999)	0.628	[0.031, 1.225]
Skogli et al. (2013)	-0.015	[-0.450, 0.420]
Riddle et al. (2013)	-0.162	[-0.504, 0.179]
Rosch et al. (2015), Seymour and al. (2016)	-0.189	[-0.609, 0.232]
Rucklidge et al. (2001), (2002)	0.569	[0.047, 1.092]
Chan et al. (2014)	-0.353	[-0.647, -0.059]
Sonuga-Barke et al. (2007)	-0.190	[-0.519, 0.138]
Wang et al. (2015)	-0.346	[-0.767, 0.076]
Wehmeier et al. (2012)	-0.266	[-0.451, -0.081]
Rucklidge et al. (2006)	0.080	[-0.473, 0.633]
Oie et al. (2016)	-0.121	[-0.581, 0.320]
Chen et al. (2008)	-0.417	[-1.064, 0.230]
Coccini et al. (2009)	0.211	[-0.841, 1.263]
Summary	**-0.081**	[**-0.187, 0.025**]



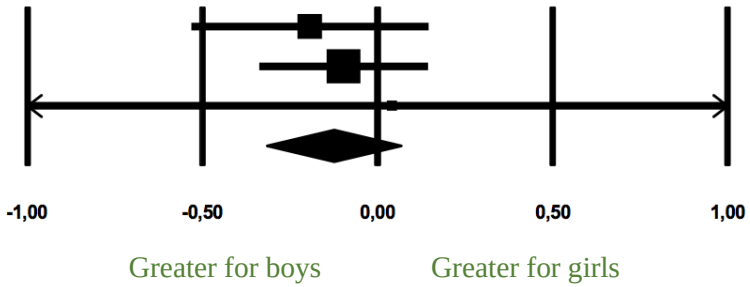
Supplementary Figure 2 Hyperactivity-impulsivity symptoms: Differences between boys and girls with ADHD
Note. CI= confidence interval. A positive effect size indicated more severe symptoms in girls, whereas a negative effect size indicated more severe symptoms in boys.

Study	Hedges'g [95% CI]		Forest plot
Borschuk et al. (2017)	-0.163	[-0.516, 0.191]	
Diaz-Orueta et al. (2014)	-0.627	[-1.220, -0.035]	
Mayfield et al. (2016)	-0.082	[-0.320, 0.156]	
Sonuga-Barke et al. (2008)	0.032	[-1.155, 1.219]	
Summary	**-0.154**	[**-0.339, 0.030**]	



Supplementary Figure 3 Hyperactivity symptoms: Differences between boys and girls with ADHD
Note. CI= confidence interval. A positive effect size indicated more severe symptoms in girls, whereas a negative effect size indicated more severe symptoms in boys.

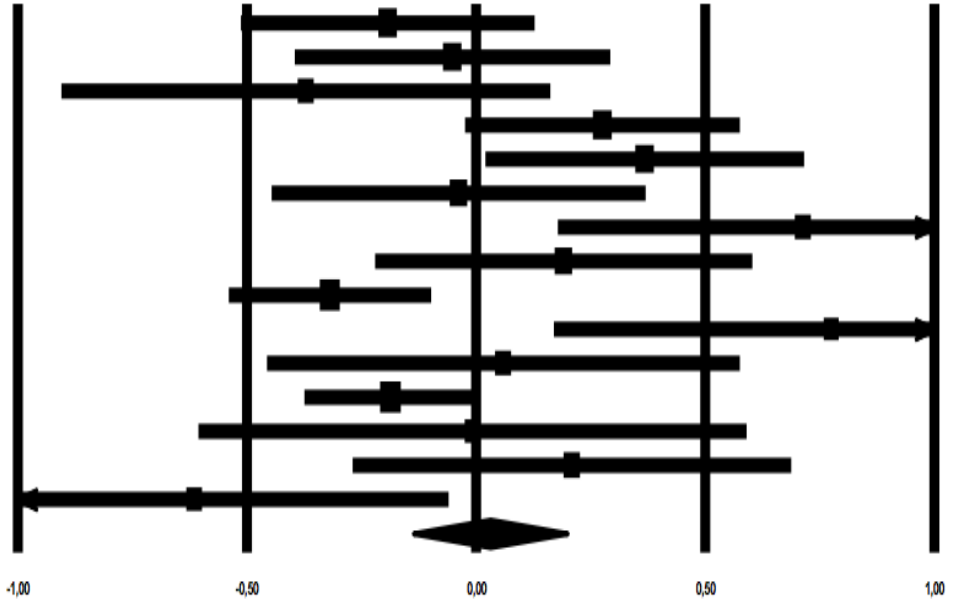
Study	Hedges'g [95% CI]		Forest plot
Gökçe et al. (2017)	-0.193	[-0.528, 0.143]	
Mayfield et al. (2016)	-0.096	[-0.334, 0.141]	
Sonuga-Barke et al. (2008)	0.042	[-1.133, 1.218]	
Summary	**-0.124**	[**-0.315, 0.068**]	



Supplementary Figure 4 Impulsivity symptoms: Differences between boys and girls with ADHD
Note. CI= confidence interval. A positive effect size indicated more severe symptoms in girls, whereas a negative effect size indicated more severe symptoms in boys.

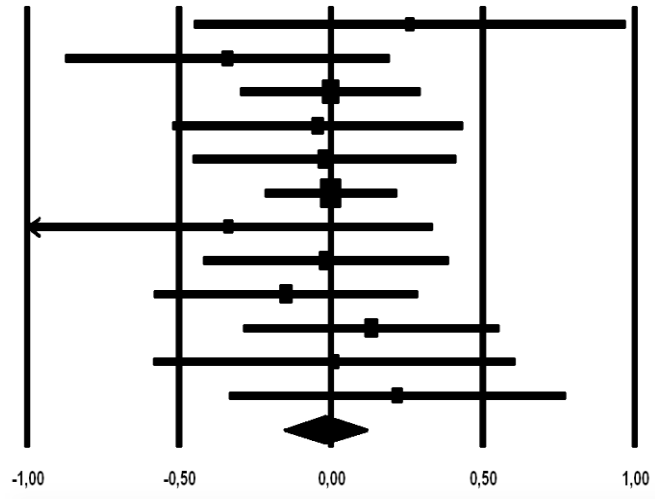
Study	Hedges'g [95% CI]	
Gökçe et al. (2017)	-0.193	[-0.510, 0.125]
Graetz et al. (2005), (2006)	-0.052	[-0.393, 0.290]
Günther et al. (2010)	-0.372	[-0.902, 0.159]
Günther et al. (2015)	0.276	[-0.021, 0.572]
Heckel et al. (2009)	0.368	[0.023, 0.713]
Isaksson et al. (2016)	-0.038	[-0.443, 0.367]
Mahone et al. (2009), O'Brien et al. (2010)	0.714	[0.182, 1.245]
Montiel-Nava et al. (2007)	0.191	[-0.218, 0.600]
Newcorn et al. (2001)	-0.319	[-0.536, -0.102]
Nolan et al. (1999)	0.0775	[0.173, 1.377]
Rucklidge et al. (2001), (2002)	0.059	[-0.453, 0.572]
Wehmeier et al. (2012)	-0.187	[-0.371, -0.002]
Yang et al. (2004)	-0.008	[-0.603, 0.587]
Zhou et al. (2006)	0.209	[-0.266, 0.685]
Derks et al. (2006)	-0.615	[-1.167, -0.063]
Summary	**0.031**	[**-0.137, 0.199**]

Forest plot



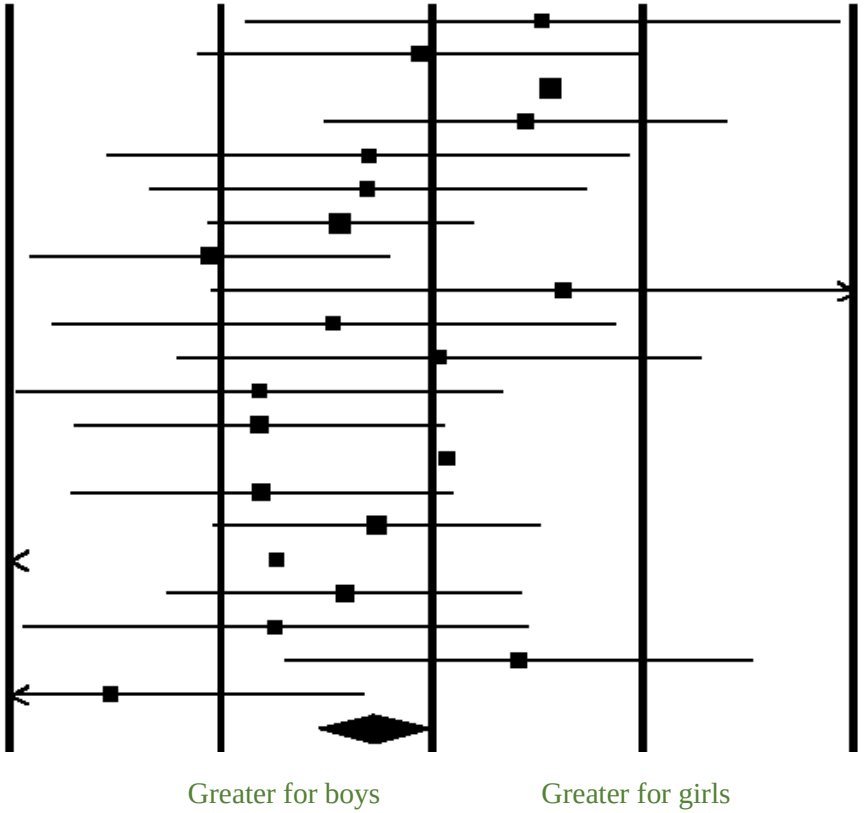
Supplementary Figure 5 Total ADHD symptoms: Differences between boys and girls with ADHD
Note. CI= confidence interval. A positive effect size indicated more severe symptoms in girls, whereas a negative effect size indicated more severe symptoms in boys.

Study	Hedges'g	[95% CI]
Alford (2007)	0.260	[-0.449, [0.969]
Günther et al. (2010)	-0.340	[-0.872, 0.191]
Günther et al. (2015)	-0.001	[-0.296, 0.294]
Huang et al. (2012)	-0.044	[-0.519, 0.432]
Lockwood et al. (2001)	-0.021	[-0.454, 0.411]
Newcorn et al. (2001)	-0.000	[-0.217, 0.216]
Wodka et al. (2007) (2008)	-0.337	[-1.006, 0.333]
Radonovich (2002)	-0.016	[-0.419, 0.386]
Rosch et al. (2015), Seymour et al. (2016)	-0.148	[-0.581, 0.285]
Wang et al. (2015)	0.133	[-0.287, 0.552]
Yang et al. (2004)	-0.011	[-0.583, 0.604]
Rucklidge (2006)	0.219	[-0.334, 0.772]
Summary	**-0.016**	**[-0.132, 0.101]**



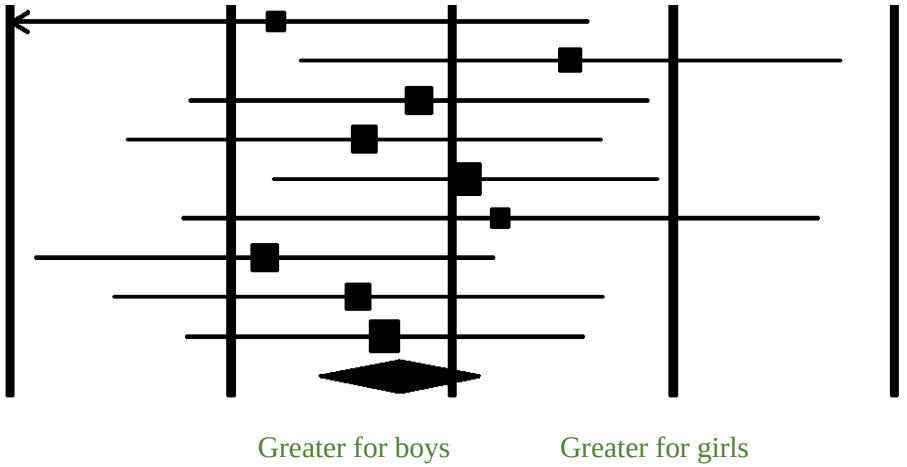
Supplementary Figure 6 Attention: Differences between boys and girls with ADHD
Note. CI= confidence interval. A positive effect size indicated a greater attention deficit in girls, whereas a negative effect size indicated a greater attention deficit in boys.

Study	Hedges'g	[95% CI]
Alford (2007)	0.261	[-0.443, 0.966]
Günther et al. (2010)	-0.029	[-0.556, 0.498]
Günther et al. (2015)	0.282	[-0.015, 0.578]
Huang et al. (2012)	0.221	[-0.256, 0.698]
Lockwood et al. (2001)	-0.151	[-0.768, 0.466]
O'Brien et al. (2010)	-0.153	[-0.670, 0.364]
Martel (2008)	-0.216	[-0.529, 0.096]
Newcorn et al. (2001)	-0.528	[-0.953, -0.102]
Nigg et al. (2002)	0.312	[-0.523, 1.146]
Wodka et al. (2007) (2008)	-0.233	[-0.899, 0.433]
Prout (200)	0.018	[-0.603, 0.638]
Radonovich (2002)	-0.410	[-0.987, 0.168]
Rosch and al. (2015), Seymour and al. (2016)	-0.409	[-0.846, 0.028]
Rucklidge (2001), (2002)	0.035	[-0.497, 0.567]
Rumsey (2004)	-0.404	[-0.856, 0.048]
Sjöwall et al. (2013)	-0.132	[-0.519, 0.256]
Sonuga-Barke et al. (2007)	-0.370	[-1.571, 0.831]
Wang et al. (2015)	-0.207	[-0.627, 0.213]
Yang et al. (2004)	-0.370	[-0.969, 0.229]
Rucklidge (2006)	-0.206	[-0.347, 0.759]
Oie et al. (2016)	-0.761	[-1.359, -0.163]
Summary	**-0.136***	**[-0.251, -0.021]**



Supplementary Figure 7 Motor inhibition: Differences between boys and girls with ADHD
Note. *p<.05 CI= confidence interval. A positive effect size indicated a greater executive deficit in girls, whereas a negative effect size indicated a greater executive deficit in boys.

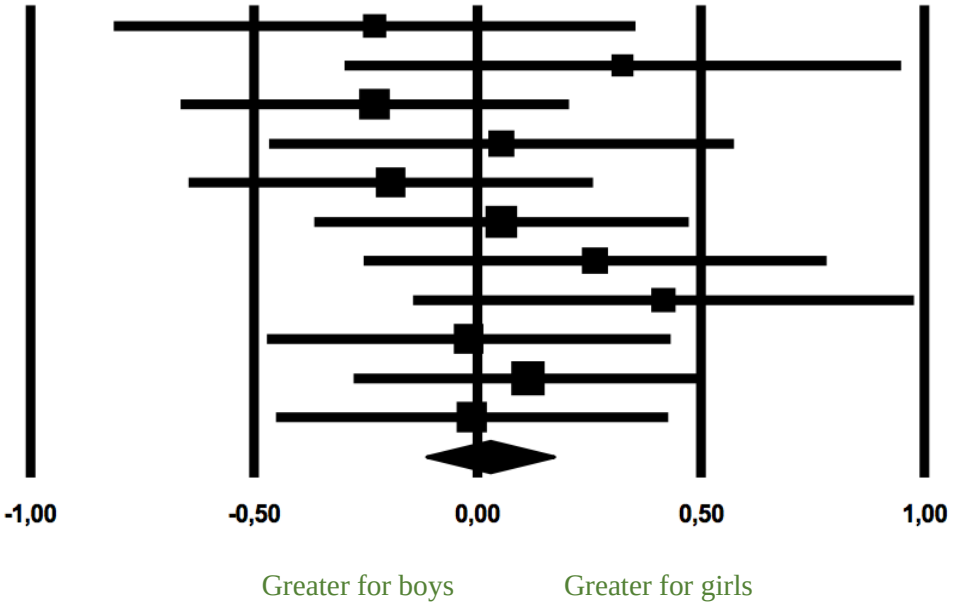
Study	Hedges'g	[95% CI]
Alford (2007)	-0.399	[-1.104, 0.307]
Lockwood et al. (2001)	0.267	[-0.344, 0.879]
Mahone et al. (2009), O'Brien et al. (2010)	-0.075	[-0.593, 0.443]
Wodka et al. (2007) (2008)	-0.199	[-0.735, 0.337]
Skogli et al. (2013)	0.030	[-0.406, 0.465]
Radonovich (2002)	0.109	[-0.611, 0.829]
Rucklidge (2001), (2002)	-0.424	[-0.942, 0.094]
Rucklidge (2006)	-0.213	[-0.766, 0.341]
Oie et al. (2016)	-0.152	[-0.601, 0.297]
Summary	**0.120**	[**-0.309, 0.069**]



Supplementary Figure 8 Interference control: Forest plot of differences between boys and girls with ADHD

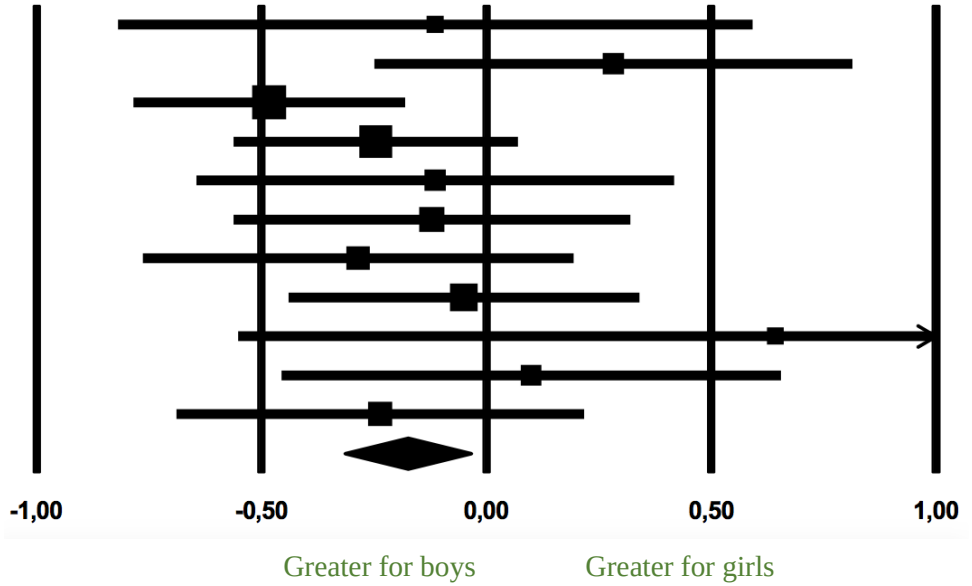
Note. CI= confidence interval. A positive effect size indicated a greater executive deficit in girls, whereas a negative effect size indicated a greater executive deficit in boys.

Study	Hedges'g	[95% CI]
Chan et al. (2016)	-0.230	[-0.811, 0.351]
Gau et al. (2009)	0.325	[-0.295, 0.944]
Lockwood et al. (2001)	-0.230	[-0.662, 0.201]
Mahone et al. (2009), O'Brien et al. (2010)	0.054	[-0.464, 0.571]
Oie et al. (2016)	-0.194	[-0.644, 0.256]
Rosch et al. (2015), Seymour et al. (2016)	0.054	[-0.362, 0.470]
Rucklidge et al. (2001), (2002)	0.263	[-0.252, 0.778]
Rucklidge (2006)	0.417	[-0.141, 0.974]
Rumsey (2004)	-0.020	[-0.468, 0.429]
Sjöwall et al. (2013)	0.113	[-0.274, 0.501]
Skogli et al. (2013)	-0.012	[-0.448, 0.424]
Summary	**0.029**	**[-0.114, 0.171]**

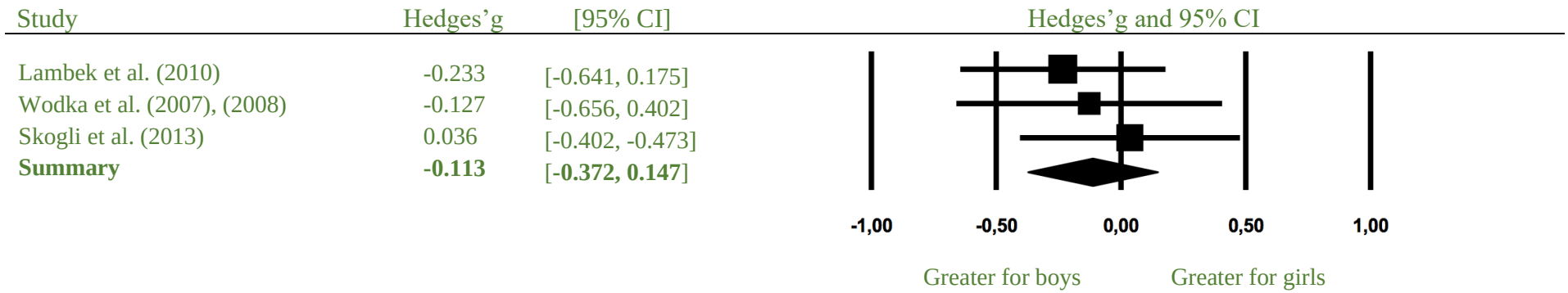


Supplementary Figure 9 Working memory: Differences between boys and girls with ADHD
Note. CI= confidence interval. A positive effect size indicated a greater executive deficit in girls, whereas a negative effect size indicated a greater executive deficit in boys.

Study	Hedges'g [95% CI]	
Alford (2007)	-0.113	[-0.815, 0.589]
Günther et al. (2010)	0.283	[-0.245, 0.812]
Günther et al. (2015)	-0.482	[-0.781, -0.183]
Martel (2008)	-0.245	[-0.558, 0.068]
Wodka et al. (2007), (2008)	-0.113	[-0.641, 0.415]
Skogli et al. (2013)	-0.120	[-0.558, 0.317]
Radonovich (2002)	-0.284	[-0.760, 0.192]
Sjöwall et al. (2013)	-0.049	[-0.436, 0.338]
Sonuga-Barke (2008)	0.644	[-0.549, 1.836]
Rucklidge (2006)	0.100	[-0.452, 0.652]
Oie et al. (2016)	-0.235	[-0.685, 0.215]
Summary	**-0.174***	[**-0.314, -0.034**]



Supplementary Figure 10 Cognitive flexibility: Differences between boys and girls with ADHD
Note. *p<.05 CI= confidence interval. A positive effect size indicated a greater executive deficit in girls, whereas a negative effect size indicated a greater executive deficit in boys.



Supplementary Figure 11 Planification: Differences between boys and girls with ADHD
Note. CI= confidence interval. A positive effect size indicated a greater executive deficit in girls, whereas a negative effect size indicated a greater executive deficit in boys.

Supplementary Table 1

lysis techniques

Risk of bias assessment

	Selection bias		Performance biais		Assessment bias		Analysis bias		Reporting bias	
	Participants recruitment	Differences between groups	Attrition	Blinding (personal)	Blinding (participants)	Relevance of assessment tools		Incomplete data	Selective reporting	Results interpretation
Alford (2007)	+	+	+	-	-	+	+	+	+	+
Bilgiç, Toker, and Uysal. (2016)	?	+	+	?	-	+	+	+	+	+
Borschuk, Rodweller, and Salorio (2018)	-	+	?	+	-	+	+	+	-	+
Chan, Lai, Luk, Hung, and Leung (2014)	-	+	+	+	-	+	+	+	+	+
Chan and Martinussen (2016)	+	+	+	-	-	+	+	?	+	+
Chen, Seipp, and Johnston (2008)	+	+	+	-	-	+	+	+	+	+
Coccini et al. (2009)	-	-	?	?	?	+	+	+	+	-
Derks, Hudziak, Dolan, Ferdinand, and Boomsma, (2006)	+	?	?	+	-	+	+	+	+	+
Díaz-Orueta et al. (2014)	-	-	+	-	-	?	+	-	-	+

DuPaul, Jitendra, and Tresco (2006)	+	+	?	+	-	-	+	+	+	+
Elkins, Malone, Keyes, Iacono, and McGue (2011)	+	+	+	+	+	+	+	+	+	+
Gau, Chiu, Shang, Cheng, and Soong (2009)	-	-	?	-	-	+	+	+	-	+
Gökçe,Yusufoğlu, Akin, and Ayaz (2017)	-	+	?	?	-	+	+	+	+	+
Graetz, Sawyer, and Baghurst (2005)	?	-	+	?	-	+	+	+	+	+
Günther, Herpertz-Dahlmaa, and Konrad (2010)	-	?	?	+	+	+	+	+	+	+
Günther, Knospe, Herpertz - Dahlman, and Konrad (2015)	-	+	+	-	-	+	+	+	+	+
Hartung et al. (2002)	+	+	?	+	-	+	+	+	+	+
Heckel, Clarke, Barry, McCarthy, and Selikowitz (2009)	-	-	+	-	-	+	+	-	+	+
Huang, Wang, and Chen (2012)	-	+	+	?	-	+	+	+	+	+
Isaksson, Ruchkin, and Lindblad (2016)	-	?	?	-	+	+	+	+	+	+
Lahey et al. (2007)	-	-	+	+	-	+	+	+	+	+
Lambek, Trillingsgaard, Kadesjo, Damm, and Thomsen (2010)	-	-	+	-	-	-	+	+	+	+
Lockwood, Marcotte, and Stern (2001)	-	+	?	-	-	+	+	+	+	+
Mahone, Mostofsky, Lasker, Zee, and Denckla, (2009)	+	+	?	?	+	+	+	+	+	+
Martel (2008)	+	-	+	-	-	+	+	+	+	+

Mayfield et al. (2016)	+	?	?	?	+	+	+	+	+	+
Montiel-Nava, Montiel-Barbero, and Peña (2007)	-	+	?	-	-	+	+	-	+	+
Muller et al. (2011a)	-	-	+	-	-	+	+	+	+	+
Muller et al. (2011b)	?	?	?	?	?	+	+	+	+	+
Newcorn et al. (2001)	?	+	?	?	?	+	+	-	+	+
Nigg, Blaskey, Huang-Pollock, and Rappley (2002)	+	+	?	+	+	+	+	-	-	+
Nolan, Volpe, Gadow, and Sprafin (1999)	-	-	?	?	-	+	+	-	-	+
O'Brien, Dowell, Mostofsky, Denckla, and Mahone (2010)	+	+	?	?	+	+	+	+	+	+
Øie, Hovik, Andersen, Czajkowski, and Skogli (2016)	-	+	+	?	?	+	+	+	+	+
Prout (2000)	+	+	+	-	-	+	+	+	+	+
Radonovich (2002)	-	?	?	?	?	+	-	+	+	+
Riddle et al. (2013)	-	-	+	+	-	+	+	+	+	+
Rosch, Dirlikov, and Mostofsky (2015)	+	+	?	?	-	+	+	+	+	+
Rucklidge (2006)	-	+	+	+	-	+	+	+	+	+
Rucklidge and Tannock (2001)	-	+	?	+	+	+	+	+	+	+
Rucklidge and Tannock (2002)	-	+	+	?	-	+	+	+	+	+
Rumsey (2004)	-	?	?	?	?	+	+	+	+	+
Seymour, Mostofsky, and Rosch (2016)	+	+	?	?	?	+	+	+	+	+
Sjöwall, Roth, Lindqvist, and Thorell (2013)	-	+	+	?	?	+	+	+	+	+

Skogli, Teicher, Andersen, Hovik, and Øie (2013)	-	+	?	-	-	+	+	+	+	+
Sonuga-Barke et al. (2007)	?	+	?	+	+	+	+	+	+	+
Sonuga-Barke and Rubia (2008)	?	-	?	?	?	+	+	+	+	-
Wang, Chen, and Huang (2015)	-	?	?	-	-	+	+	+	+	+
Wehmeier, Schacht, Escobar, Hervas, and Dickson (2012)	?	-	?	?	?	+	+	+	+	+
Wodka et al. (2007)	+	+	?	?	?	+	+	+	+	+
Wodka et al. (2008)	+	+	?	?	+	+	+	+	-	+
Yang, Jong, Chung, and Chen (2004)	-	+	+	-	-	+	+	+	+	+
Zhou, Yi, Kuang, and Fu (2006)	-	+	?	+	-	+	+	+	+	+

Note. - : High risk of bias, + : Low risk of bias, ?: Unclear risk of bias