Table 1. Quality Score Coding and Results for Included Studies

Quality metric		Studies					
		Boisvert (2012)	Gabel et al. (2013)	Grogan-Johnson et al. (2010)	Grogan-Johnson et al. (2011)	Grogan-Johnson et al. (2013)	Ruble et al. (2013)
Power		0	20	0	0	0	0
Design		8	12	16	12	16	16
Internal validity	Diagnosis	2	0	0	2	2	2
	Matching	0	1	3	3	3	2
	Treatment	2	0	0	2	2	2
	Sessions	0	0	0	2	2	2
	Method	2	0	0	0	2	2
	Targets	0	1	0	1	1	0
	Outcomes	3	0	3	3	3	0
	Blinding	0	0	0	0	2	4
	SUBTOTAL	9	2	6	13	17	14
External validity	Population	10	10	10	10	10	10
	Model	5	10	5	5	5	5
	SUBTOTAL	15	20	15	15	15	15
Reliability	IRR	10	0	0	5	0	10
	TI	0	0	0	0	10	10
	SUBTOTAL	10	0	0	5	10	20
Total Quality Score		**42**	**54**	**37**	**45**	**58**	**65**

Note. Maximum possible score for each major construct (i.e., power, design, internal validity, external validity, reliability) is 20. Studies received 0 points for a given feature if information was not provided and could not be determined. IRR = Interrater Reliability, TI = Treatment Integrity.

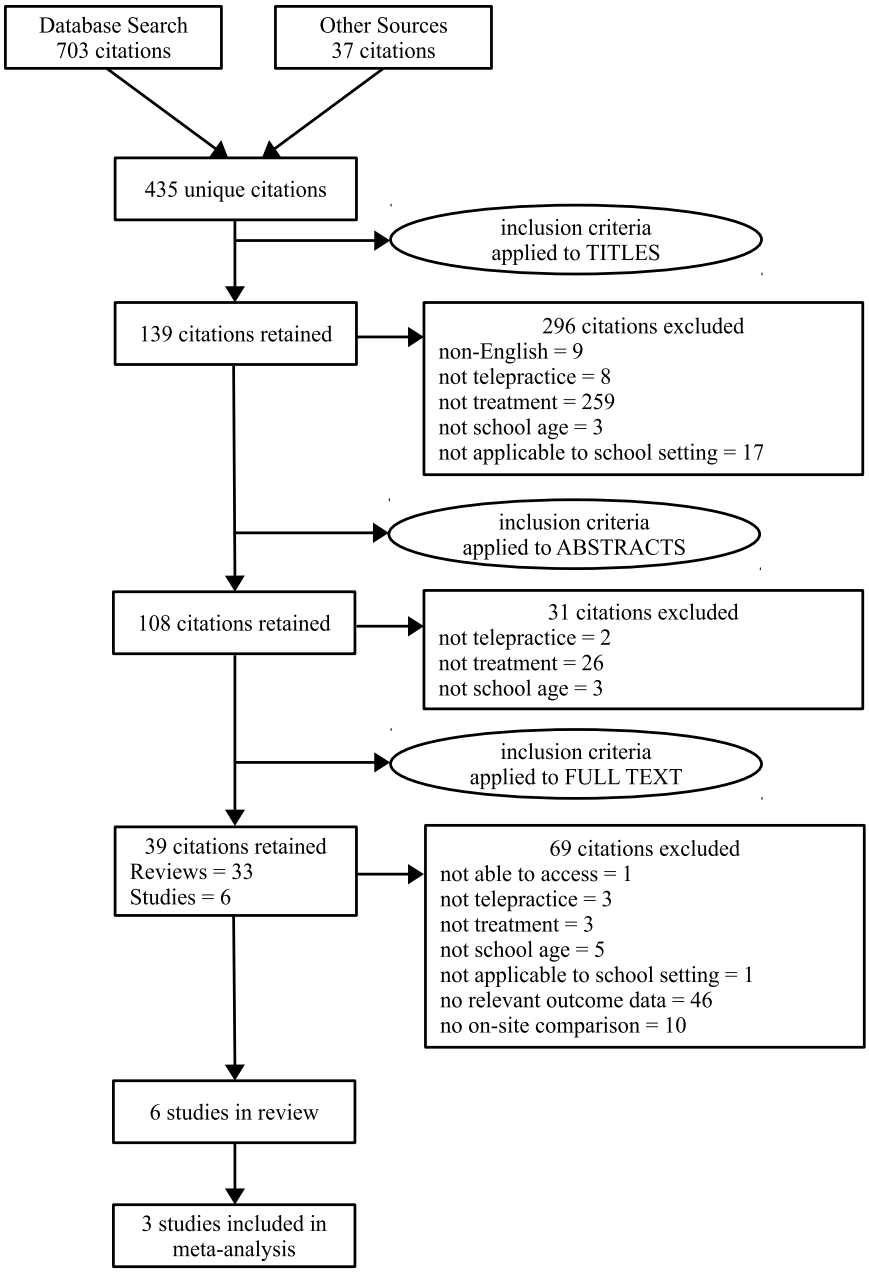

Figure 1. Results of the study search and three phase evaluation process.

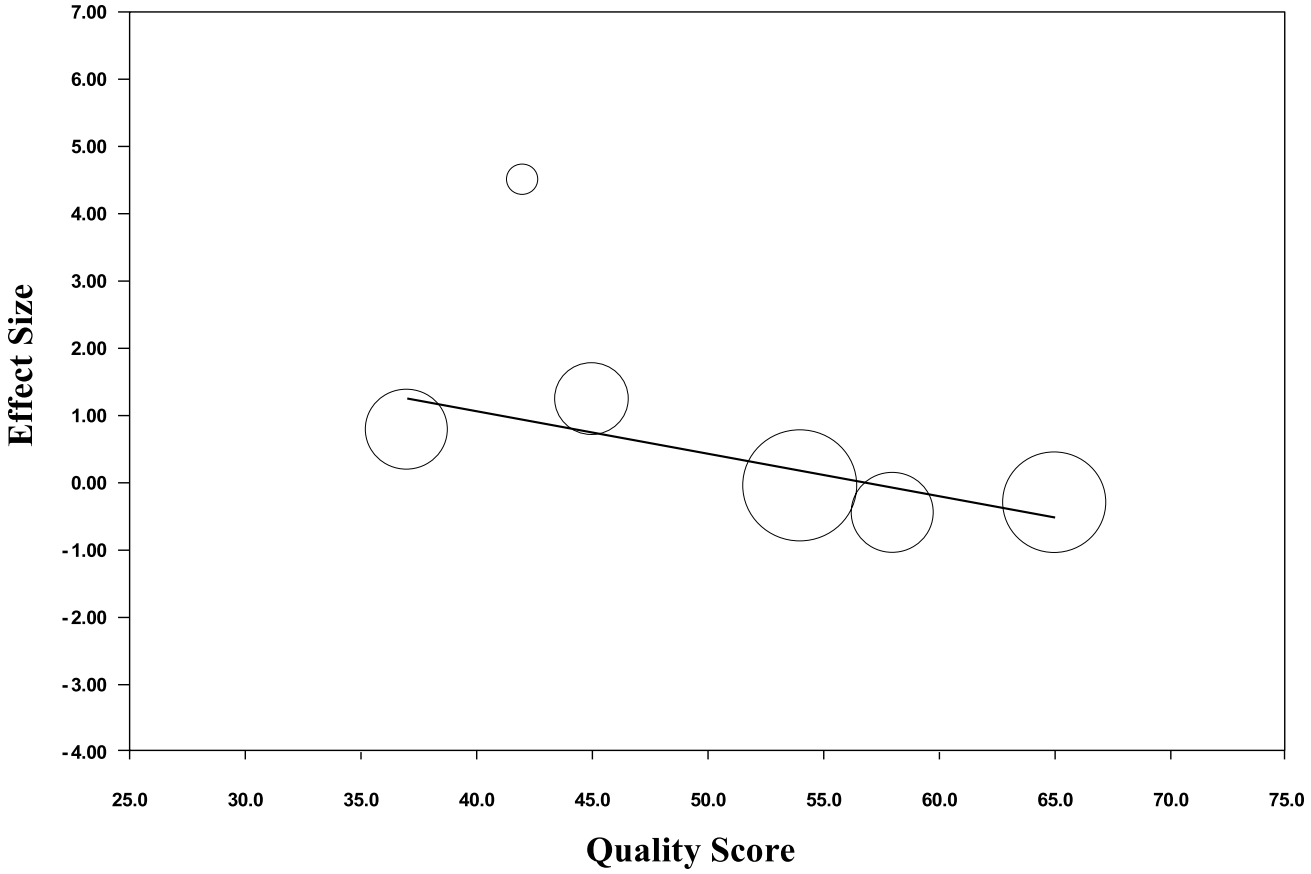

Figure 2. Results of the meta-regression analysis examining the association between study quality (quality score) and difference in improvement between telepractice and on-site groups (effect size). Quality score accounted for 56% of the variance in effect size.

Appendix A. Characteristics and Effect Sizes for Studies Included in Systematic Review

Study	**Boisvert, 2012**		
Sample Size	*N* = 6 (5 male, 1 female); 3 telepractice first, 3 on-site first		
Design	Single subject (Level 4); crossover trial		
Participants	5–12 years; diagnosed with ASD *Inclusion criteria:* must participate in mainstream class 80% of each day, communication goals on IEP, English as primary language, good health status, at least 50 words (10 oral), must meet ASHA criteria for telepractice candidacy, no other primary diagnosis, no uncorrected sensory deficits, no recent history of property destruction, no recent history of injury to self or others		
Service Delivery	30 min; 1–2x per week; 12 weeks (6 weeks per phase) Individual pull-out; school setting		
Treatment	*Intervention:* structured trials and naturalistic trials *Target (differed by student):* social engagement, answering questions, following directions, transition words, grammatical structures, speech sounds, vocabulary/concepts		
Telepractice Specifications	*Hardware* • clinician: not specified • student: computer with speakers, external webcam with embedded microphone *Software:* Skype™, Adobe® Connect Now, PresenceLearning *Connection:* not specified *Privacy:* encrypted connection to commercial servers		
Effect Size	**Outcome**	**Cohen's *d***	**95% Confidence Interval**
	Based on probe performance (unit = probe response)		
	Phase I probes (Δ from baseline)	4.50[b]	1.19–7.81
	Phase II probes (Δ from baseline)	1.99	–0.19–4.16
Study	**Gabel et al., 2013**		
Sample Size	*N* = 71 (45 male, 26 female); 71 telepractice, 5,332 on-site (database sample)		
Design	Nonrandomized controlled trial (Level 3)		
Participants	5–15 years; diagnosed with speech impairment, language impairment, learning disorder *Inclusion criteria:* no autism, no cognitive impairment, no cerebral palsy, no cleft lip/palate, no neurological impairment, no significant hearing loss, no significant visual impairment		
Service Delivery	~ 10 hours total across the academic year Individual pull-out, group pull-out, self-contained, collaborative consultative; school setting		
Treatment	*Intervention:* not specified *Target (differed by student):* intelligibility, fluency, pragmatics, speech sounds, language comprehension, language production, reading comprehension, voice		
Telepractice Specifications	*Hardware* • clinician: desktop, webcam with built-in microphone, headset • student: desktop, webcam with built-in microphone, headset • facilitator: headset *Software:* Polycom® PVX™ *Connection:* broadband internet *Privacy:* encrypted connection to commercial servers		

Appendix A. Characteristics and Effect Sizes for Studies Included in Systematic Review (continued)

Effect Size	**Outcome**	**Cohen's *d***	**95% Confidence Interval**
	Based on Functional Communication Measure improvement (unit = participant)		
	Speech sounds	0.23[a]	–0.25–0.71
	Language production	–0.35[a]	–0.81–0.11
	Language comprehension	–0.17[a]	–0.70–0.37
	Intelligibility	0.10[a]	–0.49–0.70
	Combined effect	–0.05[b,c]	–0.56–0.46
Study	**Grogan-Johnson et al., 2010**		
Sample Size	*N* = 34 (25 male, 13 female); 17 telepractice first, 17 on-site first		
Design	Randomized controlled trial (Level 2); crossover trial		
Participants	4–12 years; diagnosed with articulation disorder, language disorder, fluency disorder, learning disorder *Inclusion criteria:* current IEP, no autism or PDD, no severe cognitive deficit, no severe emotional disturbance		
Service Delivery	8 months (4 months per phase) Individual pull-out (telepractice), group pull-out (on-site); school setting		
Treatment	*Intervention:* not specified *Target (differed by student):* intelligibility, speech sounds, language production		
Telepractice Specifications	*Hardware* • clinician: computer, headphones, document camera • student: computer, headphones • facilitator: headphones *Software:* not specified *Connection:* broadband internet *Privacy:* not specified		
Effect Size	**Outcome**	**Cohen's *d***	**95% Confidence Interval**
	Based on GFTA-2 improvement (unit = participant)		
	Phase I GFTA-2 increase	0.20[a]	–0.72–1.13
	Phase II GFTA-2 increase	–0.44	–1.42–0.54
	Based on IEP goal progress (unit = goal)		
	Phase I adequate progress/mastered	1.37[a]	0.19–2.55
	Phase II adequate progress/mastered	0.18	–0.42–0.77
	Combined effect	0.79[b]	–0.27–1.84
Study	**Grogan-Johnson et al., 2011**		
Sample Size	*N* = 13 (11 male, 2 female); 7 telepractice, 6 on-site		
Design	Nonrandomized controlled trial (Level 3)		
Participants	6–11 years; diagnosed with speech sound disorder *Inclusion criteria:* communication impairment, IEP with goals for speech sound disorder, no autism or PDD, no cognitive deficits, no severe emotional disturbance, no visual impairment, no hearing impairment, no ESL students		
Service Delivery	20 min; ~2x per week; 6 months Individual pull-out; school setting		

Appendix A. Characteristics and Effect Sizes for Studies Included in Systematic Review (continued)

Treatment	*Intervention:* traditional articulation approach *Target:* speech sounds		
Telepractice Specifications	*Hardware* • clinician: desktop, webcam with built-in microphone, headset • student: desktop, webcam with built-in microphone, headset • facilitator: headset *Software:* not specified *Connection:* broadband internet *Privacy:* encrypted connection to commercial servers		
Effect Size	**Outcome**	**Cohen's *d***	**95% Confidence Interval**
	Based GFTA-2 performance (unit = probe response)		
	GFTA-2 (Δ from baseline)	2.09[a]	0.74–3.45
	Based on sound probe performance (unit = probe response)		
	Sound probe (Δ from baseline)	0.80[a]	–0.64–2.23
	Based on goal progress (unit = goal)		
	Mastered	0.83[a]	–0.01–1.67
	Combined effect	1.24[b]	0.00–2.48
Study	**Grogan-Johnson et al., 2013**		
Sample Size	*N* = 14 (9 male, 5 female); 7 telepractice, 7 on-site		
Design	Randomized controlled trial (Level 2)		
Participants	6–10 years; diagnosed with speech sound disorder *Inclusion criteria:* no significant hearing loss, no significant visual impairment, no cerebral palsy, no cognitive impairment, no cleft lip/palate, no neurological impairment, English as primary language		
Service Delivery	30 min; 2x per week; 5 weeks Individual; university clinic setting		
Treatment	*Intervention:* traditional articulation approach *Target:* speech sounds		
Telepractice Specifications	*Hardware* • clinician: desktop, webcam with built-in microphone, headset • student: laptop, webcam with built-in microphone, headset *Software:* Polycom® PVX™ *Connection:* broadband internet *Privacy:* encrypted connection to commercial servers		
Effect Size	**Outcome**	**Cohen's *d***	**95% Confidence Interval**
	Based on GFTA-2 subtest performance (unit = probe response)		
	GFTA-2 (Δ from baseline)	–0.45[b,c]	–1.51–0.61
	Based on listener judgment of sound accuracy (unit = speech sound)		
	Listener judgment (Δ from baseline)	0.45	–0.65–1.56

Appendix A. Characteristics and Effect Sizes for Studies Included in Systematic Review (continued)

Study	**Ruble et al., 2013**		
Sample Size	*N* = 49 special education teachers assigned to 49 students (42 male, 7 female); 17 telepractice, 16 on-site, 16 placebo		
Design	Randomized controlled trial (Level 2)		
Participants	3–9 years; diagnosed with ASD *Inclusion criteria:* diagnosis confirmed with ADOS, special services designated in IEP		
Service Delivery	3-hour introductory session (1x); 90 min coaching session (4x) every 5 weeks; academic year Individual coaching; school setting		
Treatment	*Intervention:* COMPASS *Target (differed by student):* communication, social skills, independence		
Telepractice Specifications	*Hardware* • clinician: not specified • student/teacher: laptop, webcam, headphones, video camera *Software:* Adobe® Connect Pro *Connection:* not specified *Privacy:* encrypted connection to a university server		
Effect Size	**Outcome**	**Cohen's *d***	**95% Confidence Interval**
	Based on PET-GAS change scores (unit = rating) [intent-to-treat analysis used]		
	PET-GAS (Δ from baseline) *Telepractice vs. On-site*	–0.30[b,c]	–0.98–0.39
	PET-GAS (Δ from baseline) *Telepractice vs. Placebo*	1.21	0.47–1.96

[a] included in combined effect

[b] included in meta-regression

[c] included in meta-analysis

Appendix B. Study Quality Rating Protocol

Quality Metric		Category	Score
Power		Inadequate	0
		Adequate	20
Design		Level 5	4
		Level 4	8
		Level 3	12
		Level 2	16
		Level 1	20
Internal Validity	Diagnosis	Different	0
		Same	2
	Group Matching	No statistical comparison, no matching reported	0
		No statistical comparison, matching reported	1
		Statistical comparison yields difference	2
		Statistical comparison yields similarity	3
	Treatment Protocol	Different	0
		Same	2
	Number of Sessions	Different	0
		Same	2
	Service Delivery Method	Different	0
		Same	2
	Treatment Targets	Different general targets	0
		Same general targets, different individual targets	1
		Same individual targets	2
	Outcome Measures	Not objective	0
		Objective	3
	Assessor Blinding	None	0
		Partial	2
		Total	4
External Validity	Study Population	Not applicable	0
		Applicable	10
	Service Delivery Model	Infeasible length/frequency	0
		Feasible length/frequency	5
		Infeasible service delivery method	0
		Feasible service delivery method	5

Appendix B. Study Quality Rating Protocol (continued)

Quality Metric		Category	Score
Reliability	Interrater Reliability	< 20% of outcome data	0
		≥ 20% of outcome data	5
		< 80% agreement	0
		≥ 80% agreement	5
	Treatment Integrity	< 20% of treatment sessions	0
		≥ 20% of treatment sessions	5
		< 80% correct implementation	0
		≥ 80% correct implementation	5
Highest Possible Quality Score			100