

---

# Intervention to Improve Attitudes Toward Stuttering: A Multi-Site International Replication and Expansion

---

[Kenneth O. St. Louis](#)<sup>\*</sup>, Ben Bolton-Grant<sup>†,††</sup>, [Autumn Cannon](#)<sup>†</sup>, Edna J. Carlo<sup>†</sup>, [Sveta Fichman](#)<sup>†</sup>, [Shweta Gupta](#)<sup>†</sup>, Krittika Kunda<sup>†</sup>, Hailey M. O'Como<sup>†</sup>, Catherine Porter<sup>†</sup>, [Bárbara M. Pratts Pérez](#)<sup>†</sup>, [Isabella Reichel](#)<sup>†,††</sup>, [Anne Z. Williams](#)<sup>†</sup>, [Salman Abdi](#)<sup>††</sup>, Elizabeth F. Aliveto<sup>††</sup>, Ann Beste-Guldborg<sup>††</sup>, Agata Błachnio<sup>††</sup>, Timothy Flynn<sup>††</sup>, [Lejla Junuzović-Žunić](#)<sup>††</sup>, [Aneta Przepiórka](#)<sup>††</sup>, Hossein Rezai<sup>††</sup>, Chelsea Roche<sup>††</sup>, [Mohyeddin Teimouri Sangani](#)<sup>††</sup>, Michael Azios<sup>†††</sup>, [Shin Ying Chu](#)<sup>†††</sup>, [Irena Polewczyk](#)<sup>†††</sup>, [Cara M. Singer](#)<sup>†††</sup>, [John A. Tetnowski](#)<sup>†††</sup>, [Janet S. Tilstra](#)<sup>†††</sup>, [Katarzyna Węsierska](#)<sup>†,††</sup>

Posted Date: 2 March 2026

doi: 10.20944/preprints202603.0046.v1

Keywords: stuttering; attitudes; POSHA-S; regression to the mean; “crossover” effect; attitude change; international



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Intervention to Improve Attitudes Toward Stuttering: A Multi-Site International Replication and Expansion

**Kenneth O. St. Louis** <sup>1,\*</sup>, **Ben Bolton-Grant** <sup>2,+</sup>, **Autumn Cannon** <sup>3,+</sup>, **Edna J. Carlo** <sup>4,+</sup>, **Sveta Fichman** <sup>5,+</sup>, **Shweta Gupta** <sup>6,+</sup>, **Krittika Kunda** <sup>7,+</sup>, **Hailey M. O'Como** <sup>8,+</sup>, **Catherine Porter** <sup>9,+</sup>, **Bárbara M. Pratts Pérez** <sup>10,+</sup>, **Isabella Reichel** <sup>11,+</sup>, **Anne Z. Williams** <sup>12,+</sup>, **Salman Abdi** <sup>13,+</sup>, **Elizabeth F. Aliveto** <sup>14,+</sup>, **Ann Beste-Guldborg** <sup>15,+</sup>, **Agata Błachnio** <sup>16,+</sup>, **Timothy Flynn** <sup>17,+</sup>, **Lejla Junuzović-Žunić** <sup>18,+</sup>, **Aneta Przepiórka** <sup>19,+</sup>, **Hossein Rezaei** <sup>20,+</sup>, **Chelsea Roche** <sup>21,+</sup>, **Mohyeddin Teimouri Sangani** <sup>22,+</sup>, **Michael Azios** <sup>23,+</sup>, **Shin Ying Chu** <sup>24,+</sup>, **Irena Polewczyk** <sup>25,+</sup>, **Cara M. Singer** <sup>26,+</sup>, **John A. Tetnowski** <sup>27,+</sup>, **Janet S. Tilstra** <sup>28,+</sup> and **Katarzyna Węsierska** <sup>29,+</sup>

- <sup>1</sup> Department of Communication Sciences & Disorders, West Virginia University, Morgantown, West Virginia, USA
- <sup>2</sup> School of Humanities and Social Sciences, Leeds Beckett University & Leeds Community Healthcare, Leeds, UK
- <sup>3</sup> Kelloggville Public Schools, Grand Rapids, Michigan, USA
- <sup>4</sup> School of Health Professions, Speech-Language Pathology Program, Medical Sciences Campus, University of Puerto Rico, San Juan, Puerto Rico, USA
- <sup>5</sup> Communication Disorders, Hadassah Academic College, Jerusalem, Isarel
- <sup>6</sup> Department of Clinical Psychology, Era University, Lucknow, Uttar Pradesh, India
- <sup>7</sup> Doctoral Student, Community Research and Action, College of Community and Public Affairs, Binghamton University (State University of New York), Vestal, New York, USA
- <sup>8</sup> Department of Communication Sciences and Disorders, Saint Cloud State University, St. Cloud, Minnesota, USA
- <sup>9</sup> School of Humanities and Social Sciences, Leeds Beckett University, Leeds, UK
- <sup>10</sup> Assistant Professor, Speech-Language Pathology Program, Universidad Interamericana de Puerto Rico, Metropolitan Campus, San Juan, Puerto Rico, USA
- <sup>11</sup> Graduate Program in Speech-Language Pathology, Touro University, New York City, New York, USA
- <sup>12</sup> Department of Communicative Disorders, University of Louisiana at Lafayette, Lafayette, LA, USA
- <sup>13</sup> Neuromuscular Rehabilitation Research Center, Research Institute of Neurosciences, Semnan University of Medical Sciences, Semnan, Iran
- <sup>14</sup> Loudoun County Public Schools, Ashburn, Virginia, USA
- <sup>15</sup> College of Education, Montana State University Billings, Billings, Montana, USA
- <sup>16</sup> Institute of Psychology, The John Paul II Catholic University of Lublin, Lublin, Poland
- <sup>17</sup> Arlington Public School, Arlington, Virginia, USA
- <sup>18</sup> Department of Speech and Language Pathology and Audiology, University of Tuzla, Faculty of Education and Rehabilitation, Tuzla, Bosnia & Herzegovina
- <sup>19</sup> The John Paul II Catholic University of Lublin, Lublin, Poland
- <sup>20</sup> Rehabilitation Research Center, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran
- <sup>21</sup> The Center at Foresight, Grand Junction, Colorado, USA
- <sup>22</sup> Department of Speech therapy, School of Paramedical and Rehabilitation of Sciences, Mashhad University of Medical Sciences, Mashhad, Iran
- <sup>23</sup> Department of Communicative Disorders, University of Louisiana at Lafayette, Lafayette, Louisiana, USA
- <sup>24</sup> Faculty of Health Sciences, Centre for Healthy Ageing and Wellness (H-CARE), Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia
- <sup>25</sup> Institute of Pedagogy, Department of Social Science, University of Silesia in Katowice, Katowice, Poland
- <sup>26</sup> Department of Communication Sciences and Disorders, Grand Valley State University, Grand Rapids, Michigan, USA

<sup>27</sup> Department of Communication Sciences and Disorders, Oklahoma State University, Stillwater, Oklahoma, USA

<sup>28</sup> Department of Communication Sciences & Disorders, St Cloud State University, St Cloud, Minnesota, USA

<sup>29</sup> University of Silesia in Katowice, Katowice, Poland

\* Correspondence: ken.stlouis@mail.wvu.edu

† These authors contributed equally to data collection of the 2026 cohort for this work.

†† These authors contributed equally to data collection of the 2016 cohort for this work.

††† These authors contributed equally to supervision in the 2026 cohort for this work.

## Abstract

**Background:** Negative public attitudes promote undesirable stereotypes and stigma in stutterers. **Method:** To mitigate negative attitudes, 16 international samples of 403 total respondents took the *Public Opinion Survey of Human Attributes–Stuttering (POSHA–S)* before and after interventions and were compared to seven combined control groups with 249 respondents. Investigators sought (a) to replicate an extreme case of regression to the mean, the “crossover” effect reported earlier in larger combined samples where negative changers with high pre-scores ended with low post-scores and positive changers with low pre-scores finished with the high post-scores and (b) to identify *POSHA–S* items related to attitudes overall change and among negative, minimal, and positive changers. **Results:** As in previous studies, stuttering attitudes improved in the intervention group but not the control group. Intervention and control respondents demonstrated “crossover” but less than the earlier samples due to lower pre-post correlations. Item contributions to pre-post change and differences among the three changer groups were inconsistent; however, high agreement items by respondents were less likely to vary than items with less unanimous agreement. **Conclusion:** The “crossover” effect was replicated, and future research should explore its presence in other measures or conditions.

**Keywords:** stuttering; attitudes; *POSHA–S*; regression to the mean; “crossover” effect; attitude change; international

---

## Introduction

It has been well established that public attitudes toward stuttering are characterized by erroneous or negative stereotypes about stuttering that affect the development, maintenance, and symptoms of stuttering as well as the attitudes and quality of life of people who stutter (e.g., Arnold et al., 2015; Barnett et al., 2005; Blood et al., 2011; Boyle & Cheyne, 2024; Bricker-Katz et al., 2013; Craig et al., 2009; Ham, 1990; Hughes, 2015; Iimura & Miyamoto, 2022; Kumar & Varghese, 2018; Norman et al., 2023; Turnbull, 2006). Weidner and St. Louis (2023) offered guidelines for developing interventions designed to reduce negative stuttering attitudes toward stuttering in children or adults who do not stutter. They summarized 47 different intervention studies that had been intended to improve stuttering attitudes. Based on authors’ objective or subjective reports, 21% of the participants had “very positive” changes, 43% had “positive” changes, 30% had “little change,” 2% had “negative and positive” change, 2% had “little and positive” change, and 2% had “negative” change. For ten studies wherein control groups with no intervention were employed, 80% showed little change and 20% improved.

Weidner and St. Louis’s (2023) list of intervention studies was not exhaustive but contained most of the available intervention studies. Included were numerous unpublished as well as published studies that administered the *Public Opinion Survey of Human Attributes–Stuttering (POSHA–S)* (St. Louis, 2011, 2022), which measures explicit public attitudes toward stuttering. Researchers around the world have been offered the option to administer the *POSHA–S* to their desired samples of people at no cost providing that they obtained human subject ethics approval at their respective institutions and shared copies of their respondent data with its author (this paper’s first author) to be included

in a large and growing *POSHA-S* database. The purpose of the database has been to empirically define what can be called “average” stuttering attitudes.

The *POSHA-S*, described in detail by St. Louis (2011, 2012, 2015, 2025), was developed as a “standard” measure of public attitudes such that samples from different areas and populations could be compared. It is psychometrically satisfactory in terms of validity, reliability, translatability, readability, and administration mode. Although periodically updated, the database recently contained 25,739 respondents from 45 countries. These include translations of the instrument into 28 different languages (St. Louis, 2025). In addition to an extensive demographic section and a general section that compares stuttering to four other “anchor” attributes (i.e., mental illness, obesity, left handedness, and intelligence) based on four items, stuttering attitudes are addressed in 39 items (35 items in the stuttering section and four from stuttering ratings in the general section). These are averaged into eight components (e.g., Cause), and these components into subscores. Two stuttering subscores, Beliefs and Self Reactions, are further averaged into an Overall Stuttering Score (OSS). Beliefs refer to what respondents know, surmise, or guess about stuttering but does not involve them specifically thinking about what they would do or where their beliefs come from. In that sense, Beliefs are external to the respondent. Self Reactions are internal in that, to rate an item, respondents must imagine themselves in a speaking situation with a stutterer or evaluate the extent or type of knowledge they bring to bear to their ratings.

An Obesity/Mental Illness subscore is also generated from general items. All ratings are converted to a standard scale from -100 to +100, and ratings for some items are inverted so that, consistently, higher scores reflect more positive (accurate, empathetic, and evidence-based) beliefs and reactions, whereas lower scores reflect more negative attitudes.

A series of studies, described below, were motivated by the need to explain why most efforts to improve stuttering attitudes in pre- versus post-test designs resulted in improved attitudes, but a substantial minority did not (Weidner & St. Louis, 2023). Three studies, in particular, that failed to improve public attitudes were especially puzzling. The first was a study in which Kuwaiti teachers did not improve in *POSHA-S* measured attitudes toward stuttering whereas a group of preservice education students did improve significantly (Abdalla & St. Louis, 2014). Another was an attempt to improve stuttering attitudes in American middle school students after viewing a well-known published film about children who stutter (Kuhn & St. Louis, 2015). The third was a large study of high school and university students in Poland who either viewed the Polish version of another well-known British film or viewed an illustrated presentation on stuttering (Węsierska et al. (2015). Mean post-intervention *POSHA-S* scores did not improve over pre-intervention scores in the respondents in these samples.

The source for the study series were samples from the *POSHA-S* database collated in 2016 that included both pre- and post-intervention data. Comprised were 29 different samples wherein an intervention occurred between the pre- and post-test and 12 samples with no interventions between the two *POSHA-S* administrations. These latter samples were carried out either as test-retest reliability samples or as control groups (C/R) in intervention studies. The four aggregate studies utilized either or both the intervention and the control/reliability (referred herein as “Control” samples).

The first aggregate study (St. Louis et al., 2020) classified the 29 intervention samples according to changes in three *POSHA-S* summary scores as “very successful” (VS), “successful” (S), “marginally successful” (MS), and “unsuccessful” (U). This was accomplished by determining whether the Beliefs subscore, Self Reactions subscore, and/or the Overall Stuttering Score (OSS) improved from pre- to post-test by more or less than 5 units. If all three improved by  $\geq 5$  units, the sample was categorized as VS; if two of three improved by that amount, it was deemed S; if only one of three improved, it was considered MS; and if none of the three improved, it was categorized as U. A total of 480 respondents (from 15 samples) were thereby assigned to the VS category, 109 respondents (from 3 samples) to the S category, 92 respondents (from 4 samples) to the MS category,

and 253 (from 7 samples) to the U category. The percentages of the intervention respondents who were thereby assigned to each category were as follows: VS = 51%, S = 12%, MS = 10%, and U = 27%.

Discriminant function analysis revealed that the four success categories were predicted partially by three characteristics of the interventions themselves but not at all by demographic characteristics of the 29 samples. The three intervention characteristics were: (a) content that was of interest to or involved the respondents (e.g., use of humor or personal experience or contact with people who stutter), (b) personal or emotional connections (e.g., feelings associated with stuttering), and (c) information about stuttering that is sufficient, but not overwhelming (e.g., showing videos of people who stutter rather than providing didactic descriptive information and explaining DOs and DON'Ts regarding interacting with people who stutter).

The second aggregate study (St. Louis, Aliveto, et al., 2024) first replicated early test-retest reliability studies of the *POSHA-S* by combining all 345 C/R respondents. Pre-test and post-test means were nearly identical, and the pre- versus post-test correlation for the overall attitude score was .79. Both metrics were indicative of satisfactory test-retest reliability, which was consistent with earlier reliability studies (Abdi et al., 2015; St. Louis et al., 2009; St. Louis, 2012). The second study's primary aim, however, was to explore characteristics of *individual respondents* from the 12 non-intervention samples. The authors observed that individual post-OSS values were substantially *negatively* correlated with the post-minus-pre scores (or amount and direction the respondents changed) indicating that something unusual may have occurred. As a next step, respondents were sorted according to whether their OSSs (a) worsened from pre- to post-test by 5 or more units ( $\leq -5$ ) (negative changers), (b) improved from pre- to post-test by 5 or more units ( $\geq +5$ ) (positive changers), or (c) did not worsen or improve by 5 units ( $> -5$  and  $< +5$ ) (minimal changers). Two surprising findings emerged, both of which were unexpected by all investigators of the 12 samples. First, rather than a large majority, only about one-third of the respondents were in the minimal change group, with the other two-thirds split quite evenly between the positive changers and negative changers. Second, confirming the correlations, the positive changers had the lowest scores on the pre-test but highest scores on the post-test, while the negative changers had the opposite pattern, highest scores at pre-test and lowest scores at post-test. This pattern was termed a "crossover" effect, which basically meant that the positive and negative changers canceled each other out in the overall pre- versus post-test means, which were virtually unchanged. The "regression to the mean" phenomenon (Barnett et al., 2005) was carefully considered because it typically influences post-test scores if pre-test scores are sorted in terms of magnitude. Importantly, regression to the mean is related to the extent to which pre- and post-test scores are correlated. There is no regression to the mean if the pre- versus post-test correlation equals  $\pm 1.0$  but complete regression to the mean if the correlation equals 0. The formula to calculate the percentage regression (movement) to the mean is  $100 \times (1 - \text{pre versus post correlation})$  (Trochim, 2026). The authors concluded that regression to the mean certainly occurred, as it always does when pre-test data are sorted from higher to lower; however, given the relatively high correlation of  $r = 0.789$  between pre and post responses, its effect was negligible.

The third aggregate study (St. Louis, Abdalla, et al., 2024) utilized both the 29 intervention samples and the 12 non-intervention samples using the same success categories of the first study (VS, S, MS, and U (St. Louis et al., 2020). In the same manner as in the second study (St. Louis, Aliveto, et al., 2024), 934 *individual* respondents from each intervention category of respondents were sorted as positive, negative, and minimal changers. The 345 non-intervention respondents were included in this third study for comparison to the intervention categories. Findings showed that all four intervention success categories, like the non-intervention category in St. Louis, Aliveto, et al. (2024), demonstrated quite similar "crossover" effects of the positive and negative changers, while minimal changers, by definition, stayed roughly the same from pre-test to post-test. Moreover, for all four success categories, the OSS values of the positive changers, like those for the non-intervention respondents in the second study, were quite dramatically lowest at pre-test and, also, dramatically highest at post-test. The opposite effect occurred for the negative changers. Importantly, the *magnitude* of the positive and negative changes was remarkably similar across the categories. What

did change as a function of success was the *percentage* of respondents in each category. For example, in the VS category, 75% were in the positive change group, 18% in the minimal change group, and 7% in the negative change group. By contrast, in the U category, 41% were positive changers, 23% were minimal changers, and 35% were negative changers. Importantly, these U category percentages were similar to the percentages in the C/R category, that is, 36% positive changers, 35% minimal changers, and 30% negative changers. As with the second study, regression to the mean was considered and found to have a measurable but minimal effect on the “crossover” effect in the VS group but virtually no effect in the other categories. Overall, the intervention sample’s OSSs, Belief, and Self Reactions improved by 10 units each.

The fourth aggregate study also considered both the 29 intervention samples and the 12 non-intervention C/R samples (St. Louis et al., 2025). Using the C/R sample as a baseline, the aim of this study was to estimate the percentages of respondents who shifted—either shifting to or shifting from—positive, minimal, or negative change groups for each of the four intervention categories. For example, if the percentage of positive changers increased from the C/R percentage of 36% to that of the very successful (VS) intervention percentage of 75%, where did the “added” approximately 40% come from? Did they come from the minimal changers, negative changers, or both? The answer is “both,” that is, 23% from the potential negative changers and 17% from the potential minimal changers. In other words, something in the VS interventions impacted both of these groups in the desired direction. Progressing in the direction from more to less successful interventions, the successful (S) interventions most likely impacted the potential negative changers by shifting 15% of them to positive changers and less than 1% to positive changers from the minimal change group. The marginally successful (MS) intervention category was most similar to the non-intervention (C/R) category. Five percent from potential negative changers shifted to positive changers and 1% shifted to minimal changers. Interventions in the unsuccessful (U) category, with an overall mean difference between pre-test and post-test OSSs near zero like the C/R category, actually reduced the potential minimal change category by 13%, with 9% shifting to the positive changers and 4% to negative changers.

To the extent that this strategy yields a valid picture of what actually occurred in the minds of the respondents subjected to interventions, the good news is that most of the intervention-induced changes were in the desired direction. Excepting a 4% shift from minimal to negative in the U category, all the shifts were toward more positive (or less negative) groups (St. Louis et al., 2025). Interventions shifted respondents from both the negative change and minimal change groups to the positive change group in the VS category. Most of the positive change gain in the S category resulted from shifts from the negative change group. All intervention-induced changes in the MS category came from the negative change group, a modest percentage to the positive change group and very small percentage to the minimal change group. The U category shifted significant percentages to both the positive and negative change groups, resulting mainly in a smaller minimal change category, compared to the non-intervention category.

One of several future lines of research recommended by authors of these aggregate studies was to explore the stability of stuttering attitudes. Two recent studies were designed to accomplish that. In the first, optometry students in the northern part of India filled out the *POSHA-S* four times in succession over three months with no intervention (Gupta, St. Louis, Rastogi, et al., 2025). Comparing the second to first administrations of positive, negative, and minimal changers, an apparent “crossover” effect was negated by regression to the mean because, unlike all but one of the 41 samples in the aggregate studies, there was virtually no correlation between pre and post scores ( $r = -0.013$ ). Responses in subsequent third and fourth administrations were chaotic. The authors concluded that the respondents and the testing environments or constraints rendered the results anomalous such that the study should be repeated with more serious and consistent respondents. It is curious that the one intervention sample in the 2016 cohort with virtually no correlation between pre and post stuttering scores (OSS  $r = 0.005$ ) were teachers from Kuwait who, unlike preservice education

students, did not improve in measured attitudes after viewing a custom video on problems children who stutter have in school (Abdalla & St. Louis, 2014).

The second study (Gupta, St. Louis, Dutt, et al., 2025) was similar to the first except it introduced three brief—but slightly different—interventions between the first and second administrations of the *POSHA-S* to clinical psychology or food and nutrition students based on OSS values of their first administration. Sorted by OSSs in the first administration, the highest one-third of respondents received a short instruction that their first impressions about stutterers were probably correct, with the hypothesis that they would not then shift negatively so much. One-third of the respondents with the lowest OSSs at the first administration were briefly instructed to consider that stuttering is really not a serious problem, hypothesizing that they would therefore respond with more positive attitudes. The middle third was told that public opinion about stuttering can improve with accurate information, hypothesizing that they would rate more positively rather than stay roughly the same. The interventions also included a PowerPoint presentation on stuttering that, unfortunately in hindsight, was different for the three one-third groups. The interventions were followed in the second administration by improved OSSs in all three groups, progressively more for the lowest third, followed by the middle third, and then by the highest third from the first administration. Degrees of success of the three different interventions were not apparent; however, overall attitudes improved substantially. Using the categories in the aforementioned aggregate studies, the overall sample improvement would be categorized very successful (VS). Next, the respondents were sorted as in the St. Louis, Abdalla, et al. (2024) and St. Louis, Aliveto, et al. (2024). When sorted this way, a notable “crossover” effect did occur but was less prominent than in the earlier aggregate studies because the regression to the mean effect was greater due to a low test-retest correlation ( $r = 0.132$ ) from the first to second administration. In third and fourth administrations, post-intervention improvements progressively decreased essentially back to the level of the first administration.

### *Purpose*

The purpose of the current study was two-fold: (a) to replicate selected findings from the aforementioned aggregate studies and (b) to extend findings from both earlier and later samples to explore which items in the *POSHA-S* are most responsible for intervention-induced changes in attitudes toward stuttering and “crossover” effect.

#### Research Questions (RQs) and Hypotheses

1. To what extent would a combined intervention group improve attitudes toward stuttering compared to a combined control group? Based on earlier data, we hypothesized that the intervention group would be approximately 10 units more positive in the post-test versus the pre-test for Beliefs, Self Reactions, and Overall Stuttering Scores and that the control group would not change from pre-test to post-test.

2. To what extent would the extreme regression past the mean or “crossover” effect characterize subgroups of intervention or control respondents who are sorted by those who improved substantially, worsened substantially, or remained nearly the same from pre-test to post-test. We hypothesized that the “crossover” effect would characterize ratings of negative changers and positive changers and that the regression to the mean phenomenon would have a negligible effect.

3. To what extent are individual *POSHA-S* items more or less responsible for improvements in stuttering attitudes and/or the “crossover” effect? We hypothesized that those items with larger effect sizes in pair-wise comparison among the three change groups would be more likely to be associated with improved attitudes or the “crossover” effect than those with smaller effect sizes.

## Materials and Methods

### Data Analysis

#### Data Sorts

Figure 1 displays the number and percent of respondents in the above-described aggregate studies, referred to here as the 2026 cohort. The 403 intervention respondents were from 16 different samples, and 249 control/reliability samples comprised seven samples (i.e., Azios et al., 2021; Bolton-Grant & Porter, 2021; Cannon et al., 2026; Carlo et al., 2025; Fichman et al., 2026; Gupta, St. Louis, Dutt, et al., 2025; Gupta, St. Louis, Rastogi, et al., 2025; Harun et al., 2022; Hearne et al., 2020; Nelson, 2020; Kral, 2025; St. Louis et al., 2018; Williams et al., 2023). The intervention samples were categorized according to success and then further divided into negative, minimal, and positive change groups. Figure 2 similarly shows the categorization and sorting of the earlier 934 intervention respondents from 29 samples and 345 respondents from 12 samples. These are herein referred to as the 2016 cohort.

Total	Intervention									Control								
Samples	16									7								
Respondents	403									249								
<b>Samples Sorted:</b> Change Pre to Post	Very Successful (VS)			Successful (S)			Marginally Successful (MS)			Unsuccessful (U)			Control (C)					
Samples	10			2			3			1			7					
Respondents	247			39			69			48			249					
Percent (%)	61.3			9.7			17.1			11.9			100.0					
<b>Individual Respondents Sorted:</b> Change Pre to Post	Negative Changers (-)			Minimal Changers (0)			Positive Changers (+)			Negative Changers (-)			Minimal Changers (0)			Positive Changers (+)		
Number	26	29	192	11	4	24	17	17	35	16	15	17	67	65	117			
Percent (%)	10.5	11.7	77.7	28.2	10.3	61.5	24.6	24.6	50.7	33.3	31.3	35.4	26.9	26.1	47.0			

**Figure 1.** Summary of samples and respondents in the 2026 cohort showing sample sorts according to success in improving attitudes toward stuttering as well as individual respondents who changed negatively, minimally, or positively.

Total	Intervention												Control / Reliability							
Samples	29												12							
Respondents	934												345							
<b>Total</b>																				
<b>Samples Sorted:</b> Change Pre to Post	Very Successful (VS)	Successful (S)			Marginally Successful (MS)			Unsuccessful (U)			Control / Reliability (C/R)									
Samples	15	3			4			7			12									
Respondents	480	109			92			253			345									
Percent (%)	51.4	11.7			9.9			27.1			100.0									
<b>Individual Respondents Sorted:</b> Change Pre to Post	Negative Changers (-)	Minimal Changers (0)	Positive Changers (+)		Negative Changers (-)	Minimal Changers (0)	Positive Changers (+)		Negative Changers (-)	Minimal Changers (0)	Positive Changers (+)		Negative Changers (-)	Minimal Changers (0)	Positive Changers (+)					
Number	35	86	359		16	37	56		22	21	49		86	55	112		103	119	123	
Percent (%)	7.3	17.9	74.8		14.7	33.9	51.4		23.9	22.8	53.3		34	21.7	44.3		29.9	34.5	35.7	

**Figure 2.** Summary of samples and respondents in the 2016 cohort showing sample sorts according to success in improving attitudes toward stuttering as well as individual respondents who changed negatively, minimally, or positively.

### Interventions and Success Categories

This replication did not compare interventions among the 16 samples. Nevertheless, Supplementary Materials 1 (Table S1) summarizes them briefly. Moreover, this study did not address differences among the three change groups within the four categories of success because the sample sizes were small for the successful, marginally successful, and unsuccessful categories. As well, the unsuccessful category was taken from only one sample.

### Analyses for Research Questions (RQs)

Supplementary Materials S2–S5 (which are in a separate Microsoft Excel file) contain tabular listings for all attitude-related items which contain data from which we addressed the research questions in this study. For RQ1, determining if the improvement in attitudes in the 2026 cohort would replicate that of the 2016 cohort, dependent *t* tests were run between the pre values and post values for the 39 POSHA–S items related to stuttering attitudes. These were run for both cohorts and separately for intervention and control groups. As been explained in numerous peer-reviewed articles (e.g., St. Louis, 2012; 2025), a Bonferroni corrected alpha level of  $p \leq 0.00417$  or  $0.05/12$  was used because it has been shown to provide an acceptable balance between not making Type I and Type II errors. Inferential statistical comparisons of the Beliefs, Self Reactions, Obesity/Mental Illness, and OSS scores were not reported in St. Louis, Abdalla et al. (2024) previous report of the 2016 between the intervention and control/reliability groups. Importantly, the proportion of statistically significant pair-wise comparisons is affected by sample size. With similar variances and mean differences, larger samples are more likely to be significantly different than smaller samples. For every *t* test, a Cohen's *d* effect size was also calculated as an index of the extent to which an item

would be more—or less—responsible for any overall changes in the summary scores. Importantly, whereas statistically significant differences are highlighted in yellow in the Supplementary materials, we used Cohen's *d* effect sizes of *all* pairwise comparisons to rank the 39 items in terms of amount of change from pre-test to post-test (see below).

For RQ2, or determining the extent of regression to the mean and a further “crossover” effect, the 2026 cohorts for intervention and control were directly compared to the 2016 cohorts. As in the St. Louis, Aliveto, et al. (2024) and St. Louis, Abdalla, et al. (2024) studies, we first sorted the respondents according to those (a) whose OSSs worsened from pre- to post-test by less than -5 units (regarded as negative changers), (b) whose OSSs improved by more than +5 units (regarded as positive changers), and (c) whose OSSs changed more than -5 units but less than +5 units (regarded as minimal changers). Next, we compared their pre-test and post-test scores for the three groups and reported the pre, post, and difference scores for all *POSHA-S* variables. To estimate regression to the mean and/or a “crossover” effect, we considered only the attitude ratings that comprised items, components, subscores, and OSSs—not demographic variables. Pre versus post results for Obesity/Mental Illness, Beliefs, Self Reactions, and OSS are shown in tables showing pre scores, actual post scores, and post scores. Additionally, OSSs are shown in line graphs that present the actual pre-test and post-test values followed by the pre-test and post-test after being corrected for regression to the mean.

RQ3 addressed heretofore unreported analyses, that is, the extent to which individual *POSHA-S* items were related to overall improvements in attitudes and/or the “crossover” effect. Analysis dealt only with the 2026 combined intervention sample as well as the 2026 control group.

Three approaches were carried out to explore RQ3 and each involved only the 39 stuttering items. First, the items were graphed according to the *POSHA-S* values on the standard -100 to +100 scale for the three pair-wise comparisons of the three change groups (negative versus minimal changers, negative versus positive changers, and minimal versus positive changers) to illustrate the degree of visually apparent variability from the three sorts. Second, Cohen's *ds* for the three pair-wise comparisons within the three change groups, i.e., negative versus minimal, negative versus positive, and minimal versus positive, were compared for all 39 items. In both pre-test and post-test, these effect sizes were rank-ordered from lowest to highest for the overall intervention and control samples plus for the three pairwise comparisons between the three change groups. The sum of the three pair-wise comparison effect sizes was taken as an index of similarity/difference of individual stuttering-related items. Third, correlations between and among the three change groups were calculated.

Figures were constructed to illustrate rankings of the 39 stuttering attitude items, separately color coded to identify of Beliefs or Self Reactions items, that occurred in identical ranks, within each of the same four quarters of the rank-ordered items, and within each of the same halves of the rank ordered items. Given 39 items, one “quarter” contained nine items and one “half” contained 19 items. A table was prepared to show frequency of occurrence in each of these rank co-occurrences.

Also, the comparisons of co-occurrence in the lower of upper halves of the ranks are shown in figures for pre-test and post-test, comparing intervention and control side-by-side. The stuttering items are classified according to whether the ranks of all three fell into lower half (“All Low”), two in the lower half and one in the higher half (“More Low”), one in the lower half and two in the higher half (“More High”), or all three in the higher half (“All High”). Finally, correlation coefficients between values for effect size differences were compared in tables.

## Results

### *Respondents*

Supplementary Materials Tables S2 and S3 for the 2026 cohort and Supplementary Materials Tables S4 and S5 for the 2016 cohort include complete demographics for the samples. Table 1 compares selected variables of the respondent characteristics. The 2026 intervention sample was four years older than the 2016 intervention sample and had an average of one more year of education. The

relative income (a weighted average of a 1-5 rating of one's income relative to [a] one's friends and family and [b] all of one's countrymen) of the 2026 intervention sample was somewhat higher than that of the 2026 control sample. Relative incomes were even for the 2016 samples. Large majorities of females (73% to 87%) participated in all the samples. From 28% to 43% reported having been married, and 15% to 38% being parents. A majority of respondents were students (55% to 88%), and from one-quarter to over one-half indicated that they were working. Self-identification and percentages knowing no one with various attributes were unremarkable except that 9% of the 2026 intervention sample reported that they were mentally ill, likely indicating that they did not fully appreciate the seriousness of mental illness.

**Table 1.** Selected demographic summary of the intervention and control samples in the 2026 and 2016 cohorts.

Variable	2026		2016	
	Intervention	Control	Intervention	Control
Age (Years)	27.3	29.8	23.3	29.2
Education (Years)	14.3	14.6	13.3	13.9
Relative Income Score (-100 to +100)	3	-7	6	6
Male	13%	20%	27%	27%
Female	87%	80%	73%	73%
I am/have been married	41%	40%	28%	43%
Parent	25%	38%	15%	34%
Student	65%	55%	88%	60%
Working	48%	57%	25%	46%
Self-Identification				
Multilingual	56%	75%	48%	46%
Intelligent	33%	33%	35%	34%
Left handed	8%	8%	8%	6%
Obese	7%	7%	5%	7%
Mentally Ill	9%	2%	2%	2%
Stuttering	2%	1%	1%	1%
No Persons Known				
Intelligent	2%	4%	1%	3%
Left handed	4%	6%	2%	3%
Obese	8%	16%	5%	7%
Mentally Ill	26%	45%	23%	21%
Stuttering	33%	38%	21%	17%

Respondents in the 2026 cohort represented eight countries or territories (USA, Puerto Rico (USA territory), UK, Poland, Israel, India, Malaysia, and New Zealand) and four different languages (English, Polish, Hebrew, and Malaysian). Those in the 2016 cohort represented seven countries (USA, UK, Poland, Bosnia and Herzegovina, Kuwait, Iran, and India). Languages included English, Polish, Bosnian-Serbian-Croatian, Persian (Farsi), and Kannada.

#### *Intervention Versus Control Group Differences (RQ1)*

Table 2 lists mean difference in three subscores and OSSs for intervention and control groups for the two cohorts. It reveals that the 2026 intervention sample demonstrated a larger improvement from pre-test to post-test more than did the 2016 intervention sample by 2-4 units, exceeding our hypothesized 10-unit positive change. Additionally, the 2026 controls, rather than essentially showing no mean improvements in 2016, had a slight but statistically non-significant improvement in the two stuttering subscores and OSS in the 2026 cohort with a mean of 4 units. Inexplicably, whereas the 2016 controls showed a non-significant 5-unit improvement in the Obesity/Mental Illness subscore, the 2026 controls did not change from pre to post overall.

**Table 2.** Pre-test, post-test, and difference (post minus pre) values for the three *POSHA-S* subscores and OSSs of the 2026 and 2016 cohorts.

	Intervention			Control		
	Pre	Post	Difference	Pre	Post	Difference
2026 Cohort						
Obesity/Mental Illness	-23	-22	1	-27	-27	0
Beliefs	38	52	14 <sup>a</sup>	32	37	5
Self Reactions	12	24	12 <sup>b</sup>	7	10	3
Overall Stuttering Score	25	37	12 <sup>c</sup>	19	23	4
2016 Cohort						
Obesity/Mental Illness	-33	-30	3	-36	-31	5
Beliefs	34	44	10 <sup>d</sup>	29	30	1
Self Reactions	1	11	10 <sup>e</sup>	-8	-8	1
Overall Stuttering Score	18	27	10 <sup>f</sup>	10	11	1

<sup>a</sup> statistically significant ( $p \leq 0.00417$ ): ( $t = 8.615$ ;  $p = <0.001$ ;  $d = 0.601$ ). <sup>b</sup> statistically significant ( $p \leq 0.00417$ ): ( $t = 7.323$ ;  $p = <0.001$ ;  $d = 0.517$ ). <sup>c</sup> statistically significant ( $p \leq 0.00417$ ): ( $t = 9.170$ ;  $p = <0.001$ ;  $d = 0.648$ ). <sup>d</sup> statistically significant ( $p \leq 0.00417$ ): ( $t = 7.133$ ;  $p = <0.001$ ;  $d = 0.330$ ). <sup>e</sup> statistically significant ( $p \leq 0.00417$ ): ( $t = 7.754$ ;  $p = <0.001$ ;  $d = 0.359$ ). <sup>f</sup> statistically significant ( $p \leq 0.00417$ ): ( $t = 8.856$ ;  $p = <0.001$ ;  $d = 0.410$ ). Note: Some apparent subtraction inconsistencies are due to rounding.

#### *Regression to the Mean and "Crossover" Effect (RQ2)*

Whenever pre-test scores in a dataset are sorted from high to low, and those scores are tracked in post-test scores, regression to the mean can occur (Trochim, 2026; Zhang & Tomblin, 2003). This means that the highest pre-test scores will "regress" or be closer to the mean in the post-test, just as the lowest pre-test scores will be higher or closer to the post-test mean. The extent to which regression to the mean occurs is a function of the correlation between the sorted pre-test and post-test scores. If the pre- and post-test scores are completely unrelated, i.e., a correlation of 0, then complete regression to the mean occurs. By contrast, if the correlation between the two scores is  $\pm 1.0$ , no regression to the mean will occur. Thus, the formula for calculating regression to the mean is 1 minus the pre- and post-test correlation. For example, if a mean pre-OSS improved to a mean post-OSS by 16 units and the correlation between the pre-post OSSs was 0.5, then half of the improvement of 16 units was due to regression to the mean. Therefore, the corrected improvement would be 8 units.

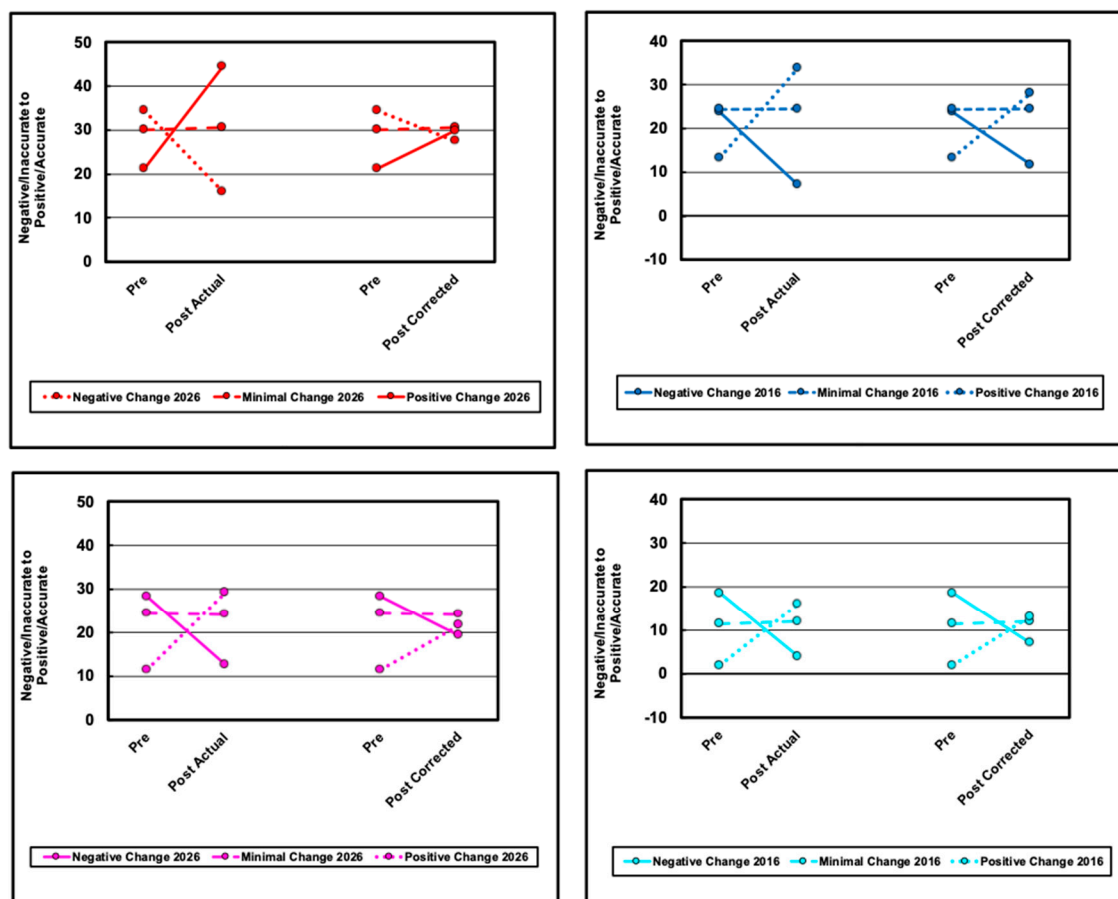
As noted in the Introduction, in the 2016 intervention group as a whole and all four success categories, as well as the control group, *POSHA-S* summary scores were characterized by not only the regression to the mean phenomenon but by a heretofore unreported extreme case of crossing the mean by nearly the same magnitude they were initially higher or lower than the mean. This, as explained earlier, was regarded as the "crossover" effect. Because the pre- versus post-test correlations were quite high (above 0.7), and the changes from pre to post were modest except for the very successful category, corrections for regression to the mean were either small or negligible.

We applied even the small corrections to the 2016 cohort and new corrections to the 2026 cohort for all attitude ratings on the *POSHA-S*. Appendix A displays the actual intervention pre, post, and difference values for Obesity/Mental illness, Beliefs, Self Reactions, and OSSs within each of the three change groups. It also shows the pre-post correlation and percent correction for regression to the mean. Appendix B shows the same information for the 2026 and 2016 cohorts for the control groups.

OSS values are shown graphically in Figure 3, which depicts the "crossover" effect for both cohorts and for intervention and control. The left lines in each graph show the actual changes from pre to post while the right lines show the post-test with a reduced effect for negative and positive changers when tempered by mathematically correcting for regression to the means. The top two graphs, for 2026 intervention (left) and control (right) show an unambiguous "crossover" effect in the actual scores. Notably, the profiles were quite similar even though the 2026 cohort's OSSs were about 10 units more positive than the 2016 cohort's OSSs. However, for the 2026 intervention sample, the

low correlation between pre and post values ( $r = 0.371$ ) diminished by nearly two-thirds the negative and positive “crossover” after applying the formula correction for regression to the mean (1- 0.371 or 62.9%). By contrast, the “crossover” effect in the 2016 cohort was much less diminished because of a much higher pre-post correlation ( $r = 0.717$ ).

Similar visual profiles also characterized the control groups in the two cohorts, as seen in the bottom two graphs of Figure 3. Yet, again, the pre- and post-OSSs of the 2026 cohort were about 10 units more positive than in the 2016 cohort. Quite similar “crossover” effects for negative and positive changers in both cohorts (-18 units for negative changers and +23 units for positive changers in 2026 and -17 units for negative changers and +20 units for positive changers in 2016) essentially cancelled out their effects for each overall control OSS means. The “crossover” effects were affected by the pre versus post correlations,  $r = 0.570$  for the 2026 controls and  $r = 0.789$  for the 2016 controls, which diminished the “crossover” effect more for the 2026 negative and positive changers than for those in the 2016 cohort. The fact that the 2026 controls improved non-significantly by 4 units (Table 2) appears to be due to the fact that the minimum changers’ means were higher than the “crossover value” than it was in the 2016 controls.



**Figure 3.** Pre- versus post-tests for the 2026 intervention group (upper left graph), 2016 intervention group (upper right graph), 2026 control group (lower left graph), and 2016 control group (lower right graph). On the left side of each graph are the actual OSSs for the three change groups (negative changers, minimal changers, and positive changers), and on the right side of each graph are the results after applying correction for regression to the mean for the negative and positive changers.

Partly confirming our hypothesis, the “crossover” effect occurred approximately equally in both intervention and control groups in the 2026 cohort. However, given the much lower pre-post correlations, the corrected post OSSs diminished it greatly. Accordingly, we can infer some yet-to-be-discovered characteristics that affect pre- to post-test correlations in the more recent samples likely have either masked or eliminated the robust “crossover” effect reported earlier (St. Louis, Abdalla, et

al., 2024; St. Louis, Aliveto, et al., 2024). Furthermore, because the pre-post correlations were markedly lower for the more recent 2026 cohort, the “crossover effect was not as large as in the 2016 cohort. In contrast, no “crossover” effect was evident for the Obesity/Mental Illness subscore.

*Stuttering Attitude Items Contributing to the “Crossover” Effect (RQ3)*

For RQ3 or exploring individual stuttering item effects on overall change from pre to post as well as differences from the three change sorts, we only considered the 2026 cohort. Figures 4–7 display the 39 stuttering attitude items clustered within the eight components for Beliefs and Self Reactions to visualize item variability as a function of respondent sorts according to negative change, minimal change, or positive change.

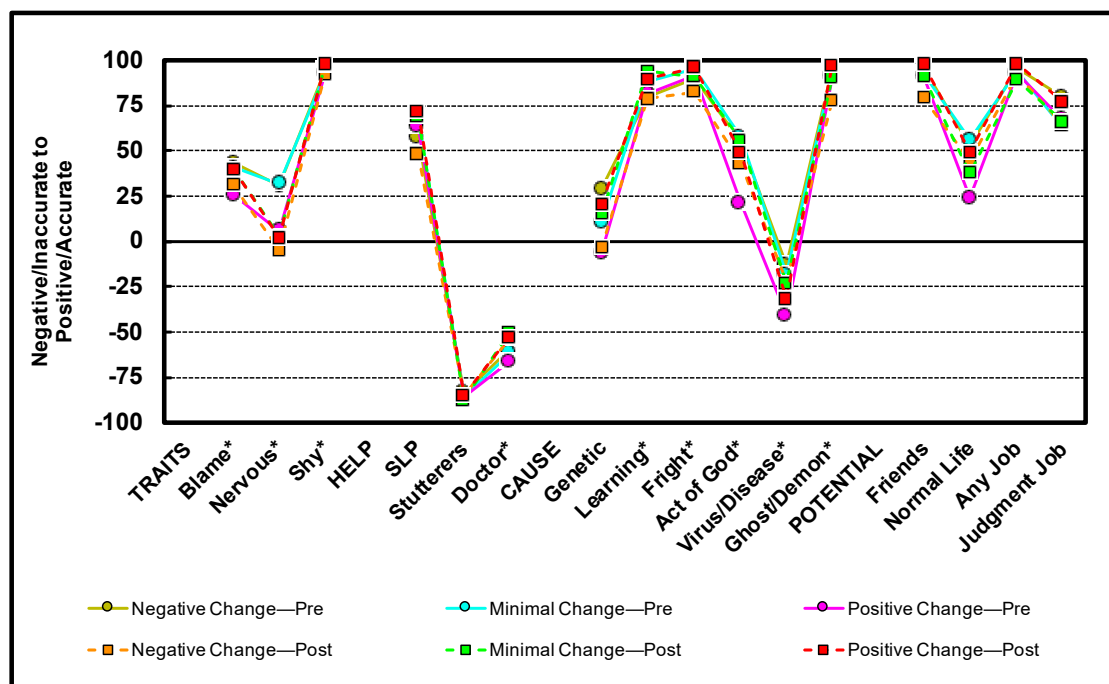
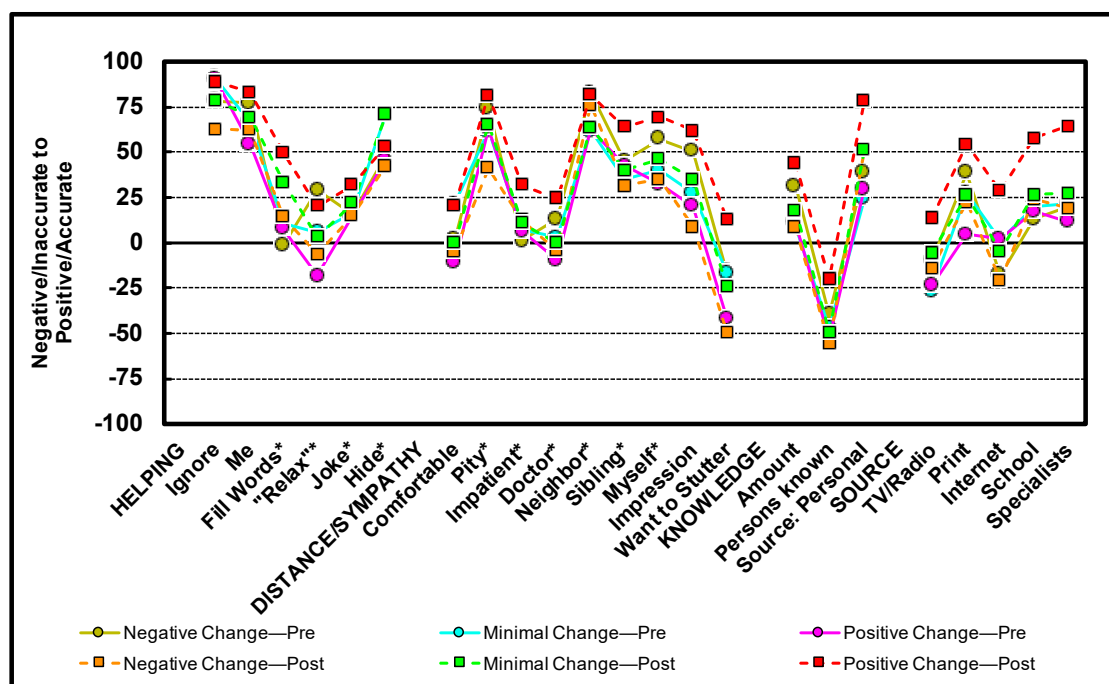
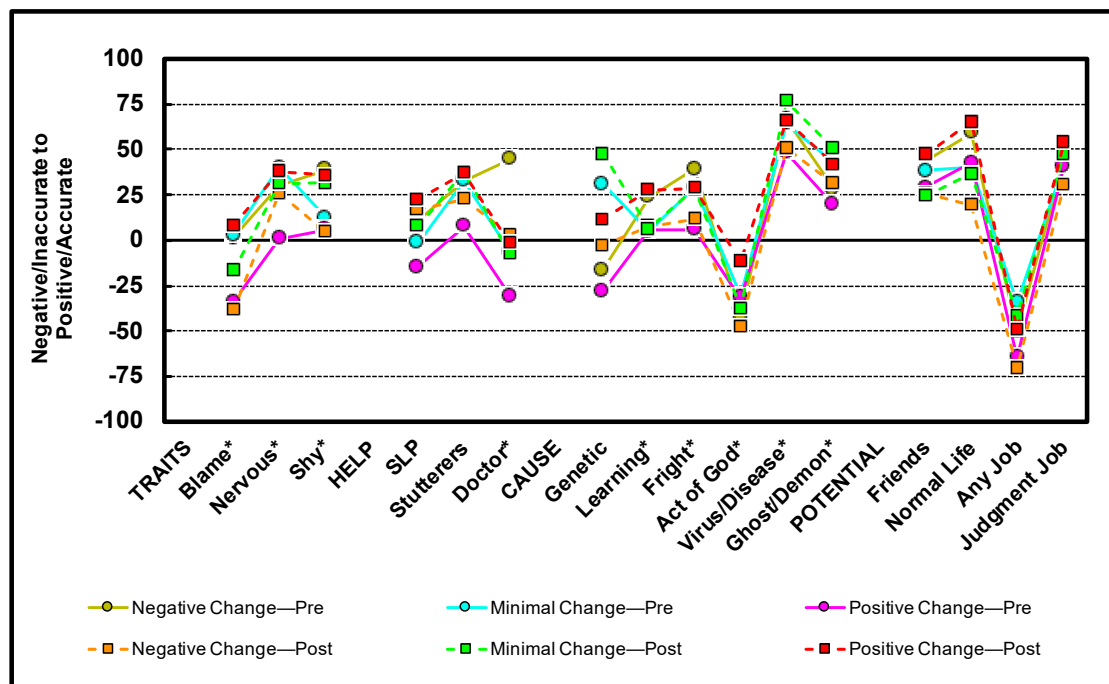


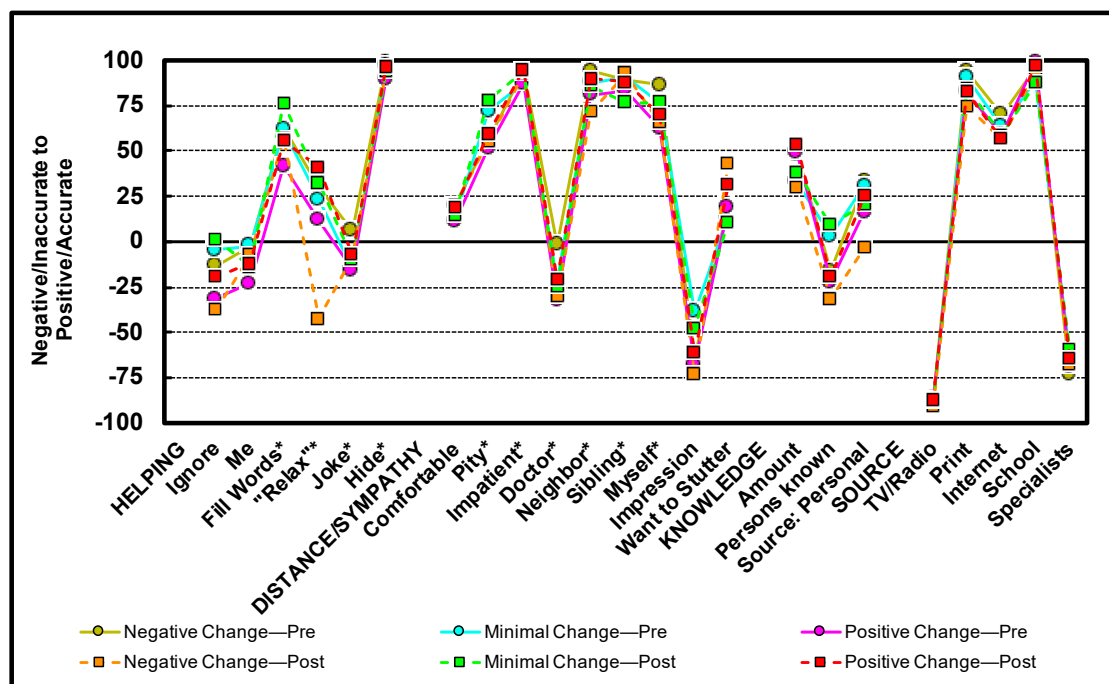
Figure 4. Stuttering item mean scores for Beliefs in the 2026 intervention sample for negative, minimal, and positive changers shown within four components.



**Figure 5.** Stuttering item mean scores for Self Reactions in the 2026 Intervention sample for negative, minimal, and positive changers shown within four components.



**Figure 6.** Stuttering item mean scores for Beliefs in the 2026 control sample for negative, minimal, and positive changers shown within four components.



**Figure 7.** Stuttering item mean scores for Self Reactions in the 2026 control sample for negative, minimal, and positive sorts shown within four components.

It is visually apparent that Self Reactions were more variable than Beliefs for the intervention sample. However, for the control sample, the reverse appears to be the case. Beliefs were more variable than Self Reactions, with the exception of advising a stuttering person to “Slow down” or “Relax” within the negative change group.

The primary RQ3 question asked was, “Can individual items that were most likely to differentiate the three sorted groups be identified?” As detailed in the Data Analysis section, they involved the magnitude of pair-wise differences (Cohen’s  $d$  values).

Supplementary Materials (Figures S6A, S6B, S7A and S7B) list the 39 stuttering items rank ordered in terms of their effect sizes in pre- versus post-test samples combined followed by rank orderings in the three pairwise comparisons between the three change groups (negative versus minimal changers, negative versus positive changers, and minimal versus positive changers). Item rankings are listed for both pre-test and post-test.

Table 3 compares the lowest to highest rank ordering of the combined pre versus post Cohens’s  $d$  values for each stuttering item to the ranks of Cohen’s  $d$ ’s for the three changer comparisons, negative versus minimal, negative versus positive, and minimal versus positive. The  $d$ ’s for these three comparisons were also rank ordered from lowest to highest. On the left, the table shows number of occurrences of identical ranks (anywhere between 1 and 39) for items whose overall pre versus post  $d$ ’s matched the ranked items from any of the three change comparisons. In the intervention sample, four items had the same rank as in one of the three ranks, and one item matched two of the three. One match occurred for the intervention post-test. Single matches occurred for only three items in the control pre-test and for two in the control post-test.

The middle of Table 3 (with light shading) shows frequency of co-occurrence within each one-fourth or quarter of the items. Given 39 items, one category contained nine items, and three contained 10 items. In this categorization, an average of approximately 13 to 18 items occurred within the same quarter across all the effect size comparison ranks. About five to six items were in one of the four categories for two comparison ranks, and only about one or two were within the same quarter for all three comparison ranks.

**Table 3.** Number of identical or similar ranks of 39 stuttering attitude items to overall pre- versus post-test Cohen’s  $d$ ’s after being sorted from lowest to highest for similarly sorted  $d$ ’s for the three pair-wise comparisons of the three change groups (negative versus minimal, negative versus positive, and minimal versus positive). Shown are (a) the numbers of identical ranks, (b) the numbers and means of approximate quarters of the items, and (c) numbers and means of approximate halves of the items.

Frequency of Occurrence in Three Sorts									
Number of Occurrences	Same Rank	Within Lowest Quarter (10)	Within 2nd Lowest Quarter (10)	Within 2nd Highest Quarter (9) <sup>a</sup>	Within Highest Quarter (10)	Mean Quarter Occurrences	Within Lower Half (20)	Within Higher Half (19) <sup>a</sup>	Mean Half Occurrences
Intervention: Cohen’s $d$ –Pre									
1	4	17	10	18	6	12.75	13	10	11.5
2	1	2	10	3	9	6.00	10	13	11.5
3	0	3	0	1	2	1.50	9	7	8.0
Intervention: Cohen’s $d$ –Post									
1	1	13	19	14	8	13.50	13	7	10.0
2	0	7	4	5	5	5.25	7	13	10.0
3	0	1	1	1	4	1.75	11	8	9.5
Control: Cohen’s $d$ –Pre									
1	3	20	11	13	8	13.00	21	3	12.0
2	0	2	5	7	11	6.25	3	21	12.0
3	0	2	3	0	0	1.25	11	4	7.5
Control: Cohen’s $d$ –Post									
1	2	12	15	15	9	12.75	16	4	10.0
2	0	3	6	6	9	6.00	4	16	10.0
3	0	4	1	0	1	1.50	12	7	9.5

<sup>a</sup> Because there were 39 items, the quarter and half designations are approximate.

The right side with darker shading in Table 3 considers co-occurrence within the lower versus upper half of the 39 items, again slightly unequal with 20 and 19 items, respectively. Item ranking in either one or two of the three pair-wise change group comparisons occurred an average of about 10

to 15 times. Items that were all within either half across the three comparisons occurred about five to 10 times.

Table 3 illustrates marked dissimilarity in ranks of stuttering attitude items of the negative, minimal, or positive changers compared to the overall pre versus post ranking in terms numerical comparisons, but specific item differences are not shown. Supplementary Materials Tables S2 and S3 display effect size differences for all stuttering items. In summary fashion, Figure 8 displays the 39 attitude items in the upper versus lower half co-occurrence comparing intervention with control results. Considering the three sorts, each stuttering attitude item received a ratings as follows: (a) "All Lower," wherein the item fell within the lower half (51%) of 39-item ranks for all three change group pair-wise comparisons, (b) "More Lower," the items falling within the lower half for two of the three comparisons and in the upper half for one item, (c) "More Higher," with two out of three falling in the upper half and one in the lower half, or (d) "All Higher," wherein all three occurring within the upper half. The figure indicates that the rankings for the pre-test intervention and control groups were visually more similar than different. Items either occurred within the same category (15/39, 38%), in an adjacent category (18/39, 46%), or in two categories distant (6/39, 16%).

Variables Sorted from Low to High based on Number (0-3) in Highest vs Lowest Half of Cohen's <i>d</i> 's in Three Changer Sorts	Intervention on 0 Lower, 3 Higher	Control 0 Lower, 3 Higher	Intervention 1 Lower, 2 Higher	Control 1 Lower, 2 Higher	Intervention 2 Lower, 1 Higher	Control 2 Lower, 1 Higher	Intervention 3 Lower, 0 Higher	Control 3 Lower, 0 Higher
<b>Pre-Administration</b>								
Speech & Language Therapist							All Lower	All Lower
Genetic							All Lower	All Lower
Reject—Ghosts, Demons, Spirits*							All Lower	All Lower
Make Friends							All Lower	All Lower
Live Normal Life							All Lower	All Lower
Reject—Make Joke*							All Lower	All Lower
Reject—I Feel Pity*				More Higher			All Lower	
Impression: Stuttering				More Higher			All Lower	
Stuttering Persons Known							All Lower	All Lower
Reject—Blame for Stuttering*				More Higher	More Lower			
Other Stutterers		All Higher			More Lower			
Reject—Medical Doctor*				More Higher	More Lower			
Reject—Act of God*				More Higher	More Lower			
Reject—Fill in Words*				More Higher	More Lower			
Reject—Advise "Slow Down" or "Relax"*				More Higher	More Lower			
Reject—Should Hide Stuttering*				More Higher	More Lower			
Reject—Concern/Worry if Doctor Stuttered*				More Higher	More Lower			
Reject—Concern/Worry if Neighbor Stuttered*				More Higher	More Lower			
Want to Stutter					More Lower	More Lower		
Reject—Nervous or Excitable*			More Higher			More Lower		
Reject—Frightening Event*			More Higher			More Lower		
Do Job Requiring Good Judgment			More Higher	More Higher				
Try to Ignore			More Higher					All Lower
I Should Help			More Higher	More Higher				
I Feel Comfortable			More Higher	More Higher				
Reject—I Feel Impatient*			More Higher					All Lower
Reject—Concern/Worry if I Stuttered*			More Higher	More Higher				
Amount Known: Stuttering			More Higher	More Higher				
Knowledge Source: Personal Experience		All Higher	More Higher					
Knowledge Source: TV/Radio			More Higher					All Lower
Knowledge Source: Internet			More Higher					All Lower
Knowledge Source: Doctors, Other Specialists			More Higher	More Higher				
Reject—Shy or Fearful*	All Higher			More Higher				
Reject—Learning or Habits*	All Higher	All Higher						
Reject—Virus or Disease*	All Higher	All Higher						
Do Any Job	All Higher			More Higher				
Reject—Concern/Worry if Sibling Stuttered*	All Higher			More Higher				
Knowledge Source: Print Media	All Higher			More Higher				
Knowledge Source: School	All Higher			More Higher				

**Figure 8.** Pre-administration POSHA-S stuttering attitude items classified according to whether their Cohen's *d* values for pairwise sorts of change groups (i.e., negative versus minimal changers, negative versus positive changers, and minimal versus positive changers) were within the lowest 51% versus highest 49% in low to high ranks for the intervention and control samples. Blue shading of the variables represent Beliefs items and tan shading represents Self Reaction items. Items are not rank ordered by Cohen's *d*'s within each cluster of "All

Lower," "More Lower," "More Higher," or "All Higher." Note: The word "reject" is not included in POSHA-S items. Scores for these items are inverted.

Figure 9 shows the results for Cohen's  $d$  rankings for the same 39 POSHA-S item in the post-test. To best visualize differences between pre-test versus post-test rankings, the order of item order is the same the intervention pre-test in Figure 8. The table shows quite clearly that the ranking of the items in the post-test were markedly different from that of the pre-test, with 14/39 (36%) within the same category, 10/39 (26%) in an adjacent category, or 15/39 (38%) two categories away.

Variables Sorted from Low to High based on Number (0-3) in Highest vs Lowest Half of Cohen's $d$ 's in Three Changer Sorts	Intervention on 0 Lower, 3 Higher	Control 0 Lower, 3 Higher	Intervention 1 Lower, 2 Higher	Control 1 Lower, 2 Higher	Intervention 2 Lower, 1 Higher	Control 2 Lower, 1 Higher	Intervention 3 Lower, 0 Higher	Control 3 Lower, 0 Higher
Post-Administration							More Lower	All Lower
Speech & Language Therapist			More Higher	More Higher				
Genetic				More Higher			All Lower	
Reject—Ghosts, Demons, Spirits*					More Lower			All Lower
Make Friends			More Higher					All Lower
Live Normal Life							All Lower	All Lower
Reject—Make Joke*		All Higher			More Lower			
Reject—I Feel Pity*	All Higher							All Lower
Impression: Stuttering							All Lower	All Lower
Stuttering Persons Known	All Higher	All Higher						
Reject—Blame for Stuttering*			More Higher			More Lower		
Other Stutterers		All Higher	More Higher					
Reject—Medical Doctor*			More Higher	More Higher				
Reject—Act of God*			More Higher	More Higher				
Reject—Fill in Words*			More Higher	More Higher				
Reject—Advise "Slow Down" or "Relax"*					More Lower			All Lower
Reject—Should Hide Stuttering*			More Higher	More Higher				
Reject—Concern/Worry if Doctor Stuttered*				More Higher			All Lower	
Reject—Concern/Worry if Neighbor Stuttered*							All Lower	All Lower
Want to Stutter	All Higher	All Higher						
Reject—Nervous or Excitable*				More Higher	More Lower			
Reject—Frightening Event*			More Higher	More Higher				
Do Job Requiring Good Judgment			More Higher	More Higher				
Try to Ignore	All Higher							All Lower
I Should Help				More Higher	More Lower			
I Feel Comfortable			More Higher					All Lower
Reject—I Feel Impatient*				More Higher			All Lower	
Reject—Concern/Worry if I Stuttered*	All Higher							All Lower
Amount Known: Stuttering							All Lower	All Lower
Knowledge Source: Personal Experience				More Higher			All Lower	
Knowledge Source: TV/Radio				More Higher			All Lower	
Knowledge Source: Internet	All Higher							All Lower
Knowledge Source: Doctors, Other Specialists		All Higher	More Higher					
Reject—Shy or Fearful*			More Higher	More Higher				
Reject—Learning or Habits*	All Higher					More Lower		
Reject—Virus or Disease*		All Higher					All Lower	
Do Any Job		All Higher			More Lower			
Reject—Concern/Worry if Sibling Stuttered*					More Lower	More Lower		
Knowledge Source: Print Media	All Higher			More Higher				
Knowledge Source: School						More Lower	All Lower	

**Figure 9.** Post-administration POSHA-S stuttering attitude items classified according to whether their Cohen's  $d$  values for pairwise sorts of change groups (i.e., negative versus minimal changers, negative versus positive changers, and minimal versus positive changers) were within the lowest 51% versus highest 49% in low to high ranks for the intervention and control samples. The order of items are shown the same as for the pre-administration to better visualize changes from pre- to post-test. Blue shading of the variables represent Beliefs items and tan shading represents Self Reaction items. Items are not rank ordered by Cohen's  $d$ 's within each cluster of "All Lower," "More Lower," "More Higher," or "All Higher." Note: The word "reject" is not included in POSHA-S items. Scores for these items are inverted.

We sought correlational confirmation of the visual similarities and dissimilarities. Table 4 shows correlations of Cohen's  $d$ 's between three pair-wise comparisons (negative versus minimal changers, negative versus positive changers, and minimal versus positive changers). The table first lists correlations for these three comparisons within each of the three sorts, which are the same regardless

of negative, minimal, or positive change sorting. Next, it shows similar correlations across the three sorts. (See table legend.) In both cases, it compares correlations for interventions and controls in the pre-tests and then for post-tests. With 39 item numbers in each case, an  $r$  of 0.415 is necessary to reach significance ( $p \leq 0.05$ ). Within each of the sorts, it is apparent that only the pre-test minimal versus positive changer effect sizes were similar enough to reach significance ( $r = 0.538$ ). For the comparison across the three sorts, none of the 18 correlations were significant. Further, although non-significant, half (9/18) of these were negative correlations. For the post-test, the lower part of Table 4 shows the negative versus minimal correlations were significant for both intervention and control samples ( $r = 0.421$  and  $0.624$ , respectively). As well, the minimal versus positive sort was significant for the intervention sample ( $r = 0.767$ ). Again, none of the correlations reached significance when comparing negative, minimal, or positive changers across the three different sorts, and ten of the 18 were negative.

**Table 4.** Correlation coefficients of Cohen's  $d$  effect size of pre-test scores (upper) and post-test scores (lower) for 39 stuttering attitude items in intervention (left) and control (right) groups for three comparisons: negative changers versus minimal changers, negative changers versus positive changers, and minimal versus positive changers. First, correlations are shown for sorts within each of the same three sorts; second, correlations are shown separately for negative, minimal, and positive changers within the three different sorts.

	Cohen's $d$ Differences Pre: Intervention			Cohen's $d$ Differences Pre: Control		
	Negative vs Minimal	Negative vs Positive	Minimal vs Positive	Negative vs Minimal	Negative vs Positive	Minimal vs Positive
Within Same Negative, Minimal or Positive Sorts (Equal in All Three)	0.376	0.006	0.538 <sup>a</sup>	0.307	0.290	0.288
	Negative Sort & Minimal Sort	Negative Sort & Positive Sort	Minimal Sort & Positive Sort	Negative Sort & Minimal Sort	Negative Sort & Positive Sort	Minimal Sort & Positive Sort
Negative Changers Within Different Sorts for Negative, Minimal and Positive Change	0.290	0.000	-0.367	-0.135	0.152	-0.079
Minimal Changers Within Different Sorts for Negative, Minimal and Positive Change	0.179	-0.303	-0.108	-0.047	0.397	0.021
Positive Changers Within Different Sorts for Negative, Minimal and Positive Change	0.124	-0.113	-0.311	0.227	-0.014	0.066
	Cohen's $d$ Differences Post: Intervention			Cohen's $d$ Differences Post: Control		
	Negative vs Minimal	Negative vs Positive	Minimal vs Positive	Negative vs Minimal	Negative vs Positive	Minimal vs Positive
Within Same Negative, Minimal or Positive Sorts (Equal in All Three)	0.421 <sup>a</sup>	0.081	0.767 <sup>a</sup>	0.624 <sup>a</sup>	0.353	0.293
	Negative Sort & Minimal Sort	Negative Sort & Positive Sort	Minimal Sort & Positive Sort	Negative Sort & Minimal Sort	Negative Sort & Positive Sort	Minimal Sort & Positive Sort
Negative Changers Within Different Sorts for Negative, Minimal and Positive Change	0.028	0.071	-0.086	-0.097	-0.293	0.184
Minimal Changers Within Different Sorts for Negative, Minimal and Positive Change	0.167	-0.023	-0.319	0.214	0.041	-0.092

Positive Changers Within Different Sorts for Negative, Minimal and Positive Change	0.346	-0.147	-0.127	0.305	-0.012	-0.190
--	-------	--------	--------	-------	--------	--------

<sup>a</sup> Statistically significant correlation. The threshold of  $r \geq 0.415$  corresponds to  $p \leq 0.05$  for these correlational analyses. This differs from the Bonferroni-corrected alpha of  $p \leq 0.00417$  used for the primary pre-post comparisons in Table 2 and the Supplementary Materials, reflecting the different inferential goals of each analysis.

Together, Figures 8 and 9, Figures S6A&B–S7A&B, and Tables 3 and 4 confirm that few obvious or consistent patterns of individual *POSHA–S* stuttering items appear to be related to sorting for negative, minimal, or positive changers from pre-test to post-test, either in the intervention sample or control sample. Nevertheless, within effect size rankings, items that were uniformly rated as among the most positive or among the most negative tended to be more stable from pre-test to post-test than items with wider inter-rater variability. These items can be seen at the top with small Cohen's  $d$ 's, or low indices of difference, in Figures 8 and 9. The most variable items occurred at the bottom of the lists. As an estimate of the least to most variant from the three tables, considering both an item's rank from lowest to highest and its frequency of occurrence, the following ten items, from less to more, were least likely to vary in the three sorts:

- Other stutterers should help
- Try to ignore stuttering
- Can do any job
- Can make friends
- Can live a normal life
- Reject ghosts, demons, or spirits as causal
- Reject feeling pity
- I want to stutter
- My impression of stuttering
- Reject an emotionally traumatic event as causal

The most likely items to vary, from more to less, were as follows:

- Reject concern/worry if my doctor stuttered
- Reject stutterer to blame for their stuttering
- My impression of stuttering
- Reject feeling impatience
- Reject virus or disease as causal
- Reject stutterers are nervous or excitable
- Reject stutterers are shy or fearful
- Reject concern/worry if I stuttered
- Knowledge source—TV/radio
- Knowledge source—doctors and other specialists

Of interest, impression of stuttering appeared in both ranks. We conclude that relatively invariant items were less likely to change substantially in repeated *POSHA–S* administrations either with no intervening interventions (control groups) or with interventions whereas relatively variant items were more likely to change. Nevertheless, we submit that what unique individual or cultural attitudes that different respondents bring to any questionnaire such as the *POSHA–S*, and how those differences are likely to be differentially affected by unique interventions supersede the influences of specific items in measuring stuttering attitude change. As such, based on these results, our RQ3 hypothesis that specific stuttering attitude items would be identifiably different in the three change groups was only weakly supported.

## Discussion

### *Summary*

This pre versus post study using the *POSHA-S* compared attitudes toward stuttering of large international samples of non-stuttering persons who either had been exposed to interventions designed to improve those attitudes or who were in non-intervention control groups. Data from a 2026 cohort consisting of 16 intervention samples and seven control samples were combined. Results from these were then compared results from a 2016 cohort, 29 intervention samples and 12 control or test-retest reliability samples, reported in four previous aggregate studies (St. Louis et al., 2020; St. Louis, Abdalla, et al., 2024; St. Louis, Aliveto, et al., 2024; St. Louis et al., 2025). Those studies documented an extreme case of regression to the mean, termed a “crossover” effect, in which negative changers from pre to post began with the highest scores and ended with the lowest scores. Conversely, positive changers began with the lowest scores and ended with the highest scores. Minimal changers, by definition, scored similarly in pre- and post-administrations. The 2026 intervention and control respondents demonstrated the “crossover” effect, similar to that of the 2016 cohort; however, because their pre- versus post-test correlations were much lower, the “crossover” effect was greatly minimized after mathematically correcting for regression to the mean. Overall, the 2026 interventions were more effective than the 2016 interventions. Aside from *POSHA-S* stuttering items that were rated similarly or dissimilarly being least or most changeable, respectively, specific questionnaire items that would predict overall improvement or who would be negative, minimal, or positive changers could not be identified. It appeared that individual differences or preferences from one respondent to another were primarily responsible for the large differences among the change groups that occurred.

### *Comparison of Interventions and Controls in Two Cohorts*

The 16 interventions applied in the 2026 cohort were clearly as effective as the 29 interventions in the 2016 cohort. Not only did they improve by the benchmark identified in previous studies (St. Louis, 2012, 2015) of an average of 10 OSS units, which was the case in the previous aggregate study (St. Louis, Abdalla et al., 2024), they improved on average by 12 units, thus confirming our RQ1 hypothesis. It should be noted that in both cohorts, Beliefs and Self Reactions improved approximately equally, indicating that the interventions affected beliefs about stuttering that are external to the respondent as well as reactions and awareness of knowledge that are internal to the respondent.

### *Regression to the Mean and the “Crossover” Effect*

The “crossover” effect clearly characterized the 2016 cohort (St. Louis, Abdalla, et al., 2024; St. Louis, Aliveto, et al., 2024), which can be regarded as an extreme case of regression to the mean. In other words, rather than high or low scorers on the pre-test scoring simply closer to the mean on the post-test, their post-test scores were far beyond the mean in the opposite direction. These findings were unexpected and potentially controversial, especially since it was observed in both intervention and control or test-retest reliability samples (Figure 3). Its effect was diminished slightly by applying the formula to correct for regression to the mean, which is a function of the pre versus post correlation.

Interestingly, upon reanalysis, the “crossover” effect occurred in the unsuccessful intervention samples that were included in the aggregate studies, mentioned in the introduction. St. Louis and Kuhn (2015) introduced a video intervention to 12 middle school respondents in a group setting without a teacher present and found that numerous students were not serious about the ratings. Several of the boys laughed when a youth in the film stuttered. A second group of 36 students watched the video with a teacher present with instructions to pay attention seriously. The first subgroup was categorized as U (unsuccessful); the second as MS (marginally successful). The

Węsierska et al. (2015) high school students and university students who were either in two intervention groups or in two control groups also demonstrated the “crossover” effect. The high school student samples who were either shown a film or a witnessed a presentation on stuttering were both classified as U, as were the university students who watched the film. The university students listening to the presentation were placed into the MS category. As in the second aggregate study (St. Louis, Aliveto, et al., 2024), the near equal OSS means of the pre-tests and post-test occurred because the minimal changers did not change while the “crossover” scores of the negative changers and positive changers cancelled each other out. This also was true of the Kuwaiti teachers in the Abdalla and St. Louis (2014) study.

One of the primary purposes of the current study was to confirm or disconfirm the “crossover” effect in entirely different aggregate samples. The effect was observed quite similarly in the 2026 intervention and control samples as well, confirming its presence. However, a much larger proportion of its effect can be considered artifact due to lower pre versus post correlations in both the intervention and control groups. We conclude that the “crossover” effect, although present in the 2026 cohorts, was not as strong as in the earlier samples, only partly supporting our RQ2 hypothesis.

How can the pre versus post correlation discrepancy between the two cohorts be explained? The reason may, in part, lie in the fact that up to 10 years intervened between collecting data for samples in the two cohorts. During that period, greater international awareness of stuttering may have affected some respondents in various samples while the remaining respondents held to previously more uniform attitudes. Importantly, the correlation discrepancy was not the result of a few outliers in the contributing samples in the cohorts. Comparing pre versus post OSS correlations among the 16 intervention samples that made up the 2026 cohort, the mean was 0.377, the median was 0.383, the minimum was -0.051, and the maximum was 0.768. These compare to means, median, minimum, and maximums of 0.586, 0.625, 0.005, and 0.867, respectively, in the 29 samples in the 2016 cohort, which, except for the minimum, were uniformly higher. The same is true of the control groups. In the 2026 cohort, the correlations of the six individual samples were: mean = 0.469, median = 0.510, minimum = -0.013, and maximum = 0.741. The same correlations, respectively, for the twelve 2016 samples were 0.618, 0.646, 0.329, and 0.815, again uniformly higher than in the 2026 cohort.

Regression to the mean is a real, but often overlooked, phenomenon (Zhang & Tomblin, 2003). To our knowledge, what the few previous studies of attitudes toward stuttering labeled as the “crossover” effect has not been reported in any other social science context. Arguably, it is an epidemiological phenomenon in that it simply describes how a sample performs within a population in repeated measures. Speech-language pathologists do not encounter regression to the mean in any tangible way with their individual clients except, perhaps, in longitudinal tracking of progress or test results. Just as a baseball player with the worst batting average for one year will likely have a better batting average in the next year, excellent performance by a client on one isolated measurement will typically be followed by some diminished performance in the next measurement.

The *POSHA-S* is designed primarily as an epidemiological measure, that is, to measure attitudes of populations through sampling (St. Louis, 2015). To that extent, it might be argued that the “crossover” effect, as an expected but extreme case of regression to the mean, is not an issue in epidemiological sampling (Smith, 2016). However, virtually all the samples in either cohort were undertaken to improve attitudes toward stuttering or to compare such samples with no intervention. The ultimate stated or unstated goal of all of this research was that more accurate or sensitive beliefs and reactions would result in more informed and empathetic interactions with those who stutter and, thereby, improve their quality of life (e.g., Arnold et al., 2015; Barnett et al., 2005; Blood et al., 2011; Boyle & Cheyne, 2024; Bricker-Katz et al., 2013; Craig et al., 2009; Ham, 1990; Hughes, 2015; Iimura & Miyamoto, 2022; Kumar & Varghese, 2018; Norman et al., 2023; Turnbull, 2006; Weidner & St. Louis, 2023). As such, we submit emphatically that it does matter that a minority of respondents exposed to interventions do, in fact, begin with quite positive attitudes and then end up with quite dramatically worse attitudes in a subsequent administration of the *POSHA-S*, and vice versa. It is also important to be aware of the robust finding that only about one-third of people—or fewer—hold stable attitudes

from test to retest after no intervention. In the 2016 controls, it was 34.5%; in the 2026 controls, it was 26.1%. The remainder split between those who change from very positive to very negative or from very negative to very positive. These low percentages indicate that the assumption that most respondents “follow the mean” in pre and post studies is simply not true for attitudes toward stuttering in POSHA-S studies. As St. Louis, Abdalla, et al. (2024) proposed, tailored interventions to target those negative, minimal, and positive changers should be undertaken. Gupta, St. Louis, Rastogi, et al. (2025) reported the first attempt to tailor interventions to high, intermediate, or low scorers, but sorting pre scores into thirds did not affect the attitude changes differentially. Only sorting as was done in the two cohorts in this study revealed differences in respondent subgroups.

Is it possible that the “crossover” effect is a characteristic of only the POSHA-S? We do not believe that it is, but confirming research needs to be done. As a sample confirmation, we explored the pre and post results of 34 speech-language pathology students who filled out a 7-point bipolar adjective (or semantic differential) scale with 25 adjective pairs, e.g., “Friendly/Unfriendly” or “Talkative/Reticent” (Reichel & St. Louis, 2004; St. Louis et al., 2014). First, the pre-test mean scores across all 25 ratings were calculated and then sorted from highest to lowest. In fact, their attitudes were less positive in the post-test than the pre-test (4.25 versus 4.06, respectively). Next, the rank-ordered means were divided as closely as possible into thirds (11, 12, and 11) and then compared to post-test scores. A “crossover” effect occurred. The overall mean difference from pre to post was -0.19, but the third with highest mean in the pre-test (4.73) had the lowest mean in the post-test (3.90) and the third with the lowest pre-test mean (3.75) had the highest post-test mean (4.23). The middle third’s pre- and post-test means were 4.30 and 4.06, respectively. The correlations between all the pre-test and post-test data was 0.506. Even so, “crossover” occurred even after correction for regression to the mean. While this post-hoc analysis was preliminary, it suggests that the “crossover” effect may be a common extreme version of regression to the mean in group attitude studies or related areas such as education.

#### *Stuttering Item Contributions to Attitude Improvement and Differences in Change Groups*

Despite exploring both the magnitude of item differences in both intervention and control samples from overall pre-test to post-test, and also among three pair-wise comparisons among the three change groups, strong likelihoods that specific items would be related to improvement or to change groups were not found. Nevertheless, the analyses revealed that items for which nearly all respondents agree on their positivity or negativity were the least likely to be different from pre to post overall or to differentially weight in the three change groups. For example, respondents in all the samples tended to be in agreement that stuttering is not caused by a fright or unseen spirits. They also agreed that they would try to ignore stuttering and not feel pity. They would believe that a person who stutters can do about any job, make friends, and live a normal life. Also, they would not think that other stutterers should help a stutterer or that they, themselves, would want to stutter.

Conversely, items that typically showed the least agreement among respondents were those items most likely to change overall in interventions or to vary among the three change groups. Respondents demonstrated considerable disagreement in the strength of their opinions about viral or disease causation or the extent to which they would be impatient or worry if their doctor stuttered. Uncertainty characterized their beliefs as to whether a stutterer is to blame for the stuttering, is nervous, or is shy or fearful. Not surprisingly, especially in the intervention sample, respondents changed ratings of their sources of knowledge about stuttering.

It cannot be overstated that these item differences cannot be assumed to always apply. For example, rejecting the notion that people who stutter are nervous or excitable or that other stutterers should help a person who stutters were invariant from pre to post in the intervention sample (Figure 4), but not so in the control group (Figure 6). This could be considered evidence that substantial differences in the attitudes of the two samples existed at the pre-test stage.

The POSHA-S was designed to tap into a wide variety of constructs relating to stuttering attitudes, which can include thoughts, feelings, and actions (St. Louis et al., 2008; St. Louis, 2025). In

standard scoring and interpretation, the stuttering items are grouped according to topic within the two subscores of Beliefs and Self Reactions. The results of our analyses for RQ3 indicates that items within both Beliefs and Self Reactions subscores were among both the least and most variant. Among the more variable Beliefs, it is striking that ratings were likely to change for several of the constructs traditionally referred to as the so-called “stuttering stereotype” (Woods & Williams, 1976). These include the stereotype that stutterers are nervous, shy, reticent, weak, and psychologically involved. Apparently, these characteristics come to mind among the public, and even professionals who treat people who stutter, when they think about stuttering. For example, Ruscello et al. (1989-90) found that when undergraduate and graduate students in speech-language pathology were asked simply to write characteristics of a “typical adult stutterer” and “typical child stutterer,” the five most frequent traits for adults were frustrated, embarrassed, angry, nervous, and self-conscious. For the hypothetical child, they were frustrated, shy, embarrassed, nervous, and anxious.

### *Strengths, Limitations, and Future Research*

Large aggregate studies such as this 2026 cohort and the earlier 2016 cohort rarely can analyze raw data from numerous unfunded international investigations. The typical alternative is to conduct meta-analysis studies. This study is not a meta-analysis; all the raw data has been archived by the first author over more than 25 years. As such, it can be best conceived of as an international, multi-site study reporting comparable results.

The main strength of our study is that it stands as a major replication and, to our best knowledge, only the second report of the unexpected finding of a “crossover” effect in respondents from a large, international dataset sorted for negative or positive change of attitudes toward stuttering in pre-post designs. On the other hand, the regression to the mean formula correction resulted in a diminished effect due to unexpected low test-retest correlations. This is not a limitation of the study but a finding that begs for further investigation to explain why respondents were apparently less stable in their responding than those samples taken a number of years earlier. Two lines for future research are thereby suggested: (a) to explore reasons for greater respondent variability and (b) to determine the extent to which the “crossover” effect occurs in other measures and for topics other than stuttering attitudes.

Inferences from our study of the 2026 cohort are limited by the fact that interventions were not at all uniform, as was the case for the 2016 cohort. Also, the 2026 intervention and control groups were unbalanced because some intervention studies lacked control groups which may explain some of the variability in the control group versus the intervention group. Nevertheless, it can be regarded as a strength that the comparison between the 2026 and 2016 cohorts documented the overall success of a wide variety of interventions but with a minority showing limited or no success.

As was called for in previous reports (e.g., St. Louis, Abdalla, et al., 2024), an especially productive line of future research would be to document the stability of changed attitudes over several successive administration of a standard measure. One attempt with a non-intervention sample was greatly limited by lack of care in responding (Gupta, St. Louis, Dutt, et al. 2025). A related study with differential interventions showed an initial improvement and “crossover” effect that appeared to weaken over two successive administrations (Gupta, St. Louis, Rastogi, et al, 2025). These studies should be replicated.

## **Conclusions**

This 2026 study replicates and extends the findings of four aggregate studies of pre- versus post-test samples designed to mitigate negative attitudes toward stuttering or to compare to non-intervention samples collected prior to 2016 (St. Louis et al., 2025; St. Louis, Abdalla, et al., 2024; St. Louis, Aliveto, et al., 2024). The 2026 cohort of intervention sample improved in *POSHA-S* measured attitudes slightly more than did the 2016 sample. The samples in both cohorts were sorted according to the degree to which mean attitudes improved. Differences between the percentages categorized as “very successful,” “successful,” “marginally successful,” or “unsuccessful” were unremarkable;

however, a 10% larger percentage of the 2026 samples were categorized as “very successful.” A heretofore unreported extreme instance of regression to the mean was noted in three of the aggregate studies and termed the “crossover” effect when the individual respondents were sorted according to whether they improved substantially, worsened substantially, or remained about the same from pre- to post-test. The positive changers began with the lowest scores and ended with the highest scores. Conversely, the negative changers started with the highest scores and ended with the lowest scores. The minimal changers, by definition, remained about the same. The same “crossover” effect was observed in the 2026 intervention and control groups. However, the effect was substantially diminished after applying the formula to correct for regression to the mean. The correction is negligible when pre versus post scores are highly correlated, as was the case in 2016 cohort. In the current 2026 cohort, the correlations for intervention and control groups were much lower, thereby diminishing the corrected “crossover” effects. Individual items were carefully analyzed through sorting them according to a Cohen’s *d* index of variability among the negative, minimal, and positive changers. Clear patterns of specific items being related to the three change group comparisons did not emerge. Nevertheless, it was observed that *POSHA-S* stuttering items with the least variability from respondent to respondent changed very little in the change groups while items with substantial inter-respondent variability did change within the groups. In summary, the search for specific items driving the “crossover” effect did not yield consistent patterns. The primary finding, that high-consensus items are most stable, suggests that individual and contextual factors, rather than item-specific content, are the primary determinants of who changes and in what direction when attitudes toward stuttering are measured.

### Appendix A. Actual and Corrected Pre, Post, and Difference Scores for Three Subscores and OSS for the 2026 and 2016 Intervention Cohorts

		Obesity/Mental Illness	Beliefs	Self Reactions	OSS
2026 Intervention: Actual Negative Changer Results	Pre	-16	47	21	34
	Post	-24	33	6	16
	Difference	-8	-14	-16	-18
2026 Intervention: Corrected Negative Changer Results Pre-Post Correlation = 0.371 Correction = 62.9%	Pre	-16	47	21	34
	Corrected Post	-20	41	15	27
	Corrected Difference	-3	-8	-10	-7
2026 Intervention: Actual Minimal Changer Results	Pre	-20	43	17	30
	Post	-18	44	17	31
	Difference	2	1	0	1
2026 Intervention: Actual Positive Changer Results	Pre	-25	34	8	21
	Post	-22	59	30	44
	Difference	4	25	22	23
2026 Intervention: Corrected Positive Changer Results Pre-Post Correlation = 0.371 Correction = 62.9%	Pre	-25	34	8	21
	Corrected Post	-23	45	17	30
	Corrected Difference	1	14	13	9
2016 Intervention: Actual Negative Changer Results	Pre	-32	39	8	24
	Post	-34	20	-6	7
	Difference	-2	-20	-14	-17
2016 Intervention: Corrected Negative Changer Results Pre-Post Correlation = 0.717 Correction = 28.3%	Pre	-32	39	8	24
	Corrected Post	-33	27	-1	12
	Corrected Difference	-1	-7	-4	-12
2016 Intervention: Actual Minimal Changer Results	Pre	-32	42	6	24
	Post	-30	41	8	24
	Difference	2	-1	2	0
	Pre	-34	30	-3	13

2016 Intervention: Actual	Post	-30	52	16	34
Positive Changer Results	Difference	4	22	19	20
2016 Intervention: Corrected	Pre	-34	30	-3	13
Positive Changer Results	Corrected Post	-31	44	10	28
Pre-Post Correlation = 0.717 Correction = 28.3%	Corrected Difference	1	8	6	15

## Appendix B. Actual and Corrected Pre, Post, and Difference Scores for Three Subscores and OSS for the 2026 and 2016 Control Cohorts

		Obesity/Mental Illness	Beliefs	Self Reactions	OSS
2026 Control: Actual Negative Changer Results	Pre	-27	38	18	28
	Post	-27	22	3	13
	Difference	-1	-16	-15	-16
2026 Control: Corrected Negative Changer Results Pre-Post Correlation = 0.570 Correction = 43.0%	Pre	-27	38	18	28
	Corrected Post	-27	31	9	19
	Corrected Difference	0	-9	-6	-9
2026 Control: Actual Minimal Changer Results	Pre	-28	37	12	24
	Post	-29	36	12	24
	Difference	0	0	0	0
2026 Control: Actual Positive Changer Results	Pre	-26	25	-3	11
	Post	-26	46	12	29
	Difference	0	21	15	18
2026 Control: Corrected Positive Changer Results Pre-Post Correlation = 0.570 Correction = 43.0%	Pre	-26	25	-3	11
	Corrected Post	-26	34	6	22
	Corrected Difference	0	12	6	10
2016 Control: Actual Negative Changer Results	Pre	-32	37	0	19
	Post	-28	21	-13	4
	Difference	4	-16	-13	-15
2016 Control: Corrected Negative Changer Results Pre-Post Correlation = 0.789 Correction = 21.1%	Pre	-32	37	0	19
	Corrected Post	-28	26	-10	7
	Corrected Difference	4	-4	-3	-12
2016 Control: Actual Minimal Changer Results	Pre	-36	32	-10	11
	Post	-34	32	-8	12
	Difference	2	0	1	1
2016 Control: Actual Positive Changer Results	Pre	-38	18	-14	2
	Post	-31	34	-2	16
	Difference	8	16	12	14
2016 Control: Corrected Positive Changer Results Pre-Post Correlation = 0.789 Correction = 21.1%	Pre	-38	18	-14	2
	Corrected Post	-34	30	-5	13
	Corrected Difference	3	4	3	11

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

**Author Contributions:** Conceptualization, KOSL; Methodology, KOSL; Software, KOSL; Validation, KOSL; Formal Analysis, KOSL; Investigation, BB-G, EJC, SF, SG, KK, HMO, CP, BMPP, IR, AZW, SA, EFA, AB-G, AB, TF, LJ-Ž, AP, HR, CR, MTS, KW; Resources, n/a; Data Curation, KOSL; Writing – Original Draft Preparation, KOSL; Writing – Review & Editing, KOSL, HR, LJ-Ž; Visualization, KOSL; Supervision, MA, SYC, IP, CMS, JT, JST; Project Administration, KOSL; Funding Acquisition, n/a.

**Funding:** This research received no external funding.

**Data Availability Statement:** The dataset from which the data for this study were taken is contained in the Supplementary Materials, Tables S2–S5, and are available to future investigators.

**Acknowledgments:** We gratefully acknowledge the contributions to data collection of Fauzia Abdalla, Brendan Carr, Jocine Gloria Chandrabose, Sarah Eisert, Sheryl Gottwald, Brianne Hanlon, Hariyani Harun, Jessica Hartley, Chelsea Heaster, Kailey Holcombe, Joanna Holmes, Daniel Hudock, Laura Gibson, Kia Johnson, Ahmad Poomohammad, Mariswamy Pushpavathi, Claire Rowland, Sarah Spears, Kira Stork, and Mercedes Ware. We also thank the human subject ethics committees, administrators, and supervisors in the various institutions who facilitated the studies involved in the 64 different samples of the 2016 and 2026 cohorts combined in this report.

**Conflicts of Interest:** The authors declare no conflict of interest. The first author owns the copyright of the POSHA-S.

## References

- Abdalla, F., & St. Louis, K. O. (2014). Modifying attitudes of Arab school teachers toward stuttering. *Language, Speech, and Hearing Services in the Schools, 45*, 14-25. [https://doi.org/10.1044/2013\\_LSHSS-13-0012](https://doi.org/10.1044/2013_LSHSS-13-0012)
- Abdi, S., Jalaei, Sh., Teimouri Sangani, M., & Pourmohammad, A. (2015). مقیاس سنجش نگر □ عمومی نسبت به □ از گ ایهای انسانیت “ بومی سازی □ بررسی □ ای □ پایایی نسخه فارسی [Development and evaluation of psychometric properties of the Persian version of “public opinion survey of human attributes-stuttering”]. *Journal of Modern Rehabilitation Research Faculty of Rehabilitation, Tehran University of Medical Sciences, 9*(5), 86-98. Special Issue No. 3. [In Persian]
- Arnold, H. S., Li, J., & Goltl, K. (2015). Beliefs of teachers versus non-teachers about people who stutter. *Journal of Fluency Disorders, 43*, 28–39. <https://doi.org/10.1016/j.jfludis.2014.12.001>
- Azios, M., Kunda, K., Irani, F., & St. Louis, K. O. (2021, November). *Changing university students' attitudes about stuttering: A social way of thinking*. Poster presented at the Annual Convention of the American Speech-Language-Hearing Association, Washington, DC.
- Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology, 34*, 215–220. <https://doi.org/10.1093/ije/dyh299>
- Blood, G. W., Blood, I. M., Tramontana, G. M., Sylvia, A. J., Boyle, M. P., & Motzko, G. R. (2011). Self-reported experience of bullying of students who stutter: relations with life satisfaction, life orientation, and self-esteem. *Perceptual and Motor Skills, 113*(2), 353–364. <https://doi.org/10.2466/07.10.15.17.PMS.113.5.353-364>
- Bolton-Grant, B., & Porter, C. (2021). *Effect of an eLearning module on improving attitudes of non-traditional students*. Department of Speech and Language Therapy, Leeds Beckett University, Leeds, UK
- Boyle, M. P., & Cheyne, M. R. (2024). Major discrimination due to stuttering and its association with quality of life. *Journal of Fluency Disorders, 80*, 106051. <https://doi.org/10.1016/j.jfludis.2024.106051>
- Bricker-Katz, G., Lincoln, M., & Cumming, S. (2013). Stuttering and work life: An interpretative phenomenological analysis. *Journal of Fluency Disorders, 38*(4), 342-355. <https://doi.org/10.1016/j.jfludis.2013.08.001>
- Cannon, A., Shattuck, A., Otieno, S., Hoban, A., St. Louis, K.O., and Singer, C.M. (2025). *Exploration of documentaries in changing attitudes toward stuttering among university student* [Unpublished manuscript]. Communication Sciences and Disorders Department, Grand Valley State University.
- Carlo, E. J., Pratts-Pérez, B. M., & St. Louis, K. O. (2025). Attitudes toward stuttering of speech-language pathology students from Puerto Rico before and after completing the degree. *Journal of Fluency Disorders, 86*, 106161. <https://doi.org/10.1016/j.jfludis.2025.106161>
- Cohen, J. (1988). *Statistical power for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Craig, A., Blumgart, E., & Tran, Y. (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal of Fluency Disorders, 34*(2), 61–71. <https://doi.org/10.1016/j.jfludis.2009.05.002>
- Fichman, S, Adelman, C., Apelboim-Dushnitzky, G., Horesh, V., Israel, S., Maor, A., & St. Louis, K. O. (under review). *International Journal of Inclusive Education*.
- Gupta, S., St. Louis, K. O., Dutt, K., & Rastogi, A. (2025). Stability in attitudes of university students towards stuttering in the Indian state of Uttar Pradesh. *Journal of the All India Institute of Speech and Hearing, 44*(2), 1-14. [https://doi.org/10.4103/jose.JOSE\\_12\\_24](https://doi.org/10.4103/jose.JOSE_12_24)

- Gupta, S., St. Louis, K. O., Rastogi, A., & Dutt, K. (2025, Preprint). Effectiveness and stability of differential interventions in reducing negative attitudes toward stuttering. *Social Science Research Network (SSRN)*. <https://ssrn.com/abstract=5400504> or <http://dx.doi.org/10.2139/ssrn.5400504>
- Ham, R. E. (1990). What is stuttering: Variations and stereotypes. *Journal of Fluency Disorders*, 15, 259-273. [https://doi.org/10.1016/0094-730X\(90\)90041-P](https://doi.org/10.1016/0094-730X(90)90041-P)
- Hearne, A., Miles, A., Douglas, J., Carr, B., Nicholls, J. R., Bullock, M. S., Pang, V., & Southwood, H. (2020). Exploring teachers' attitudes: knowledge and classroom strategies for children who stutter in New Zealand. *Speech, Language and Hearing*, <https://doi.org/10.1080/2050571X.2020.1750756>
- Harun, H., Chu, S. Y., & binti Mohd. Ali, M. (2022, March). *Stuttering education for teachers: Changing perception, knowledge and attitudes*. Paper presented at the 1st National Speech Therapy Symposium. Kuala Lumpur, Malaysia.
- Hughes, S. (2015). Attitudes toward stuttering: An annotated bibliography. In K. O. St. Louis (Ed.), *Stuttering meets stereotype, stigma, and discrimination: An overview of attitude research* (pp. 310-350). Morgantown, WV: West Virginia University Press.
- Iimura, D., & Miyamoto, S. (2022). The influence of stuttering and co-occurring disorders on job difficulties among adults who stutter. *Speech, Language and Hearing*, 25(2), 235-244. <https://doi.org/10.1080/2050571X.2020.1852494>
- Kral, G. (2025). *Changing adolescents' attitudes toward stuttering as a result of an educational intervention: A quasi-experimental study*. Unpublished Master's Thesis, University of Silesia, Katowice, Poland.
- Kuhn, C. D. & St. Louis, K. O. (2015, November). *Attitudes toward stuttering of middle school students before & after a stuttering video*. Poster presented at the Annual Convention of the American Speech-Language-Hearing Association. Denver, CO.
- Kumar, A. S., & Varghese, A. L. (2018). A study to assess awareness and attitudes of teachers towards primary school children with stuttering in Dakshina Kannada District, India. *Journal of Clinical & Diagnostic Research*, 12(9).
- Nelson, H. (2020). *Changing college student attitudes toward people who stutter*. Masters thesis in Communication Sciences and Disorders. 6. St Cloud State University, St Cloud, MN [https://repository.stcloudstate.edu/csd\\_etds/6](https://repository.stcloudstate.edu/csd_etds/6)
- Norman, A., Lowe, R., Onslow, M., O'Brian, S., Packman, A., Menzies, R., & Schroeder, L. (2023). Cost of illness and health-related quality of life for stuttering: Two systematic reviews. *Journal of Speech, Language, and Hearing Research*, 66(11), 4414-4431. [https://doi.org/10.1044/2023\\_JSLHR-23-00072](https://doi.org/10.1044/2023_JSLHR-23-00072)
- Reichel, I., & St. Louis, K. O. (2004). The effects of emotional intelligence training in graduate fluency disorders classes. In H-G Bosshardt, J. S. Yaruss, & H. F. M. Peters (Eds.), *Fluency disorders: Theory, research, treatment, and self-help* (pp. 474-481). Nijmegen, the Netherlands: International Fluency Association/Nijmegen University Press.
- Ruscello, D. M., Lass, N. J., French, R., & Channel, M. D. (1989-90). Speech language pathology students' perceptions of stutterers. *National Student Speech Language Hearing Association Journal*, 17, 86-89. [https://doi.org/10.1044/nsshla\\_17\\_86](https://doi.org/10.1044/nsshla_17_86)
- Smith, G. (2016). A fallacy that will not die. *The Journal of Investing*. 25, 7-15. <https://doi.org/10.3905/joi.2016.25.1.007>
- St. Louis, K. O. (2011). The Public Opinion Survey of Human Attributes-Stuttering (POSHA-S): Summary framework and empirical comparisons. *Journal of Fluency Disorders*, 36(4), 256-261. <https://doi.org/10.1016/j.jfludis.2011.02.003>
- St. Louis, K. O. (2012). Research and development on a public attitude instrument for stuttering. *Journal of Communication Disorders*, 45(2), 129-146. <https://doi.org/10.1016/j.jcomdis.2011.12.001>
- St. Louis, K. O. (2015). Epidemiology of public attitudes toward stuttering. In K. O. St. Louis (Ed.), *Stuttering meets stereotype, stigma, and discrimination: An overview of attitude research* (pp. 7-42). Morgantown, WV: West Virginia University Press.
- St. Louis, K. O. (2022). *Public Opinion Survey of Human Attributes-Stuttering (POSHA-S)*. Morgantown, WV: Populore.

- St. Louis, K. O. (2025). An international database of public attitudes toward stuttering. *Data*, 10(9), 147. <https://doi.org/10.3390/data10090147>
- St. Louis, K. O., Abdalla, F., Abdi, S., Aliveto, E., Beste-Guldborg, A., Błachnio, A., Bolton-Grant, B., Eisert, S., Flynn, T., Gottwald, S., Hartley, J., Hudock, D., Johnson, K. N., Junuzović-Žunić, L., Przepiórka, L., Pushpavathi, M., Reichel, I., Rezai, H., Roche, C., Spears, S., Teimouri Sangani, M., & Węsierska, K. (2024). Profiles of public attitude change regarding stuttering. *Language and Health*. <https://doi.org/10.1016/j.laheal.2024.08.001>
- St. Louis, K. O., Aliveto, E. F., Teymouri Sangani, M., Abdi, S., Rezai, H., Abdalla, F., Przepiórka, A., Błachnio, A., Węsierska, K., Junuzović-Žunić, L., Eisert, S., Roche, C. K., Reichel, I., Beste-Guldborg, A., Flynn, T., Bolton, B., Gottwald, S., Spears, S., Hudock, D., Hartley, J., Pushpavathi, M., & Johnson, K. N. (2024). Measuring public attitudes toward stuttering: Test-retest reliability revisited. *Clinical Archives of Communication Disorders*. <https://doi.org/10.21849/cacd.2023.01081>
- St. Louis, K. O., Lubker, B. B., Yaruss, J. S., Adkins, T. A., & Pill, J. C. (2008). Development of a prototype questionnaire to survey public attitudes toward stuttering: Principles and methodologies in the first prototype. *The Internet Journal of Epidemiology*, 5(2). <http://ispub.com/IJE/5/2/7561>
- St. Louis, K. O., Lubker, B. B., Yaruss, J. S., & Aliveto, E. F. (2009). Development of a prototype questionnaire to survey public attitudes toward stuttering: Reliability of the second prototype. *Contemporary Issues in Communication Sciences and Disorders*, 36, 101-107. <https://doi.org/10.1016/j.jfludis.2009.02.001>
- St. Louis, K. O., Węsierska, K., & Polewczyk, I. (2018). Improving Polish stuttering attitudes: An experimental study of teachers and university students. *American Journal of Speech-Language Pathology*, 27, 1195-1210. [https://doi.org/10.1044/2018\\_AJSLP-ODC11-17-0179](https://doi.org/10.1044/2018_AJSLP-ODC11-17-0179)
- St. Louis, K. O., Węsierska, K., Przepiórka, A., Błachnio, A., Beucher, C., Abdalla, F., Flynn, T., Reichel, I., Beste-Guldborg, A., Junuzović-Žunić, L., Gottwald, S., Hartley, J., Eisert, S., Johnson, K., Bolton, B., Teimouri Sangani, M., Rezai, H., Abdi, S., Pushpavathi, M., Hudock, D., Spears, S., & Aliveto, E. (2020). Success in changing stuttering attitudes: A retrospective study of 29 intervention samples. *Journal of Communication Disorders*, 84, 1-18. 105972. <https://doi.org/10.1016/j.jcomdis.2019.105972>
- St. Louis, K. O., Węsierska, K., Reichel, I., Roche, C., Rezai, H., Abdalla, F., Junuzović-Žunić, L., Przepiórka, A., Flynn, T., Aliveto, E., Beste-Guldborg, A., Błachnio, A., Teimouri Sangani, M., Bolton-Grant, B., & Abdi, S. (2025). Modifying stuttering attitudes: Who changes and in what direction? *Forum Lingwistyczne*. 13(2), 1-18. <https://doi.org/10.31261/FL.2025.13.2.17>
- St. Louis, K. O., Williams, M. J., Ware, M. B., Guendouzi, J., & Reichel, I. (2014). The *Public Opinion Survey of Human Attributes–Stuttering (POSHA–S)* and *Bipolar Adjective Scale (BAS)*: Aspects of validity. *Journal of Communication Disorders*, 50, 36-50. <https://doi.org/10.1016/j.jcomdis.2014.04.001>
- Trochim, W. M. K. (2026). Research methods knowledge base. *Regression to the mean*. <https://conjointly.com/kb/regression-to-the-mean/>
- Turnbull, J. (2006). Promoting greater understanding in peers of children who stammer. *Emotional and Behavioural Difficulties*, 11(4), 237–247. <https://doi.org/10.1080/13632750601022139>
- Weidner, M. & St. Louis, K. O. (2023). Changing public attitudes toward stuttering. In H. Sønsterud, & K. Węsierska (Eds.) *Dialogue without barriers – A comprehensive speech therapy intervention in stuttering* (English Version). Chorzów, Poland: Agere Aude Foundation for Knowledge and Social Dialogue. <https://www.logolab.edu.pl/dialogue-without-barriers-a-comprehensive-approach-to-dealing-with-stuttering-english-version/>
- Węsierska, K., Błachnio, A., Przepiórka, A., & St. Louis, K. O. (2015, October). *Zmiana postaw wobec jąkania w Polsce – wstępne doniesienia z badań [Changing attitudes towards stuttering in Poland: Preliminary study report]*. Paper presented at the International Conference on Speech-Language Therapy: Modern Trends in Logopaedic Diagnosis and Therapy. Chorzów, Poland.
- Williams, A. Z., Tetnowski, J., & St. Louis, K. O. (2023). Addressing attitudes about stuttering in preservice teachers. *Perspectives of the ASHA Special Interest Groups*, 8, 372-379. [http://doi.org/10.1044/2023\\_PERSP-22-00187](http://doi.org/10.1044/2023_PERSP-22-00187)

Woods, C. L., & Williams, D. E. (1976). Traits attributed to stuttering and normally fluent males. *Journal of Speech and Hearing Research*, 19(2), 267–278. <https://doi.org/10.1044/jshr.1902.267>

Zhang, X, & Tomblin, J. B. (2004). Explaining and controlling regression to the mean in longitudinal research designs. *Journal of Speech, Language, and Hearing Research*, 46, 1340-1351. [https://doi.org/10.1044/1092-4388\(2003/104\)](https://doi.org/10.1044/1092-4388(2003/104))

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.