# Australian Randomised Control Trial
## *Think Equal Early Childhood Program*

**Federation University**

**Yale** Center for Emotional Intelligence

**THiNK EQUAL**

## Prepared by

Dr Susan Emmett
Dr Lynne Reeder
Dr Chris Turville
Karen Schneider
**Federation University Australia**

Assistant Professor Prof Craig Bailey
**Yale University**

**August 2021**

**About Think Equal**

The Think Equal program is a concrete, step by step social and emotional learning program for children, which has now been implemented in early childhood education settings in several countries including Singapore, Canada, and Botswana. Indeed, programs are currently running in 14 additional countries on five continents.

Think Equal lessons are delivered by early childhood educators and are organized into 36 topic areas containing evidence-based resources that emphasize narratives in literature and in the wider context of children's lives. In addition, the Think Equal curricular focuses on using positive, accurate language as well as social cognition and learning within a context of social justice.

**Australian evaluation**

During 2019 researchers at Federation University Australia partnered with the globally recognized Yale University to determine any changes in children's social and emotional behaviour after participating in the Think Equal program. All educators delivering Think Equal underwent a seven-hour training session based on the Think Equal training manual. A national training session was held for lead teachers in Brisbane and Melbourne. Educators and teachers that took part in these sessions went on to train other early childhood teachers who delivered the program in centres. This mentoring and training ensured that the program was consistently understood and could be translated into practice using the specific learning materials. The implementation of the thirty-week program was overseen by members of the Australian Think Equal team throughout 2019.

**Methodology**

The Think Equal evaluation in Australia adopted a cluster randomised controlled trial design; these experimental designs are considered the "gold standard" for exploring program impact[1]. Unlike correlational designs, randomized controlled trials explore causality by comparing the treatment group to a counterfactual, comparable group (i.e., the control group). Randomization reduces the likelihood of allocation bias and increases the likelihood of baseline equivalence between treatment and controlled groups. Centres were randomly assigned to either the treatment or control group from within the Australian States of Victoria and Queensland. In addition, 7 children were randomly selected for data collection in each classroom. Forty-five centres ($n_{\text{Think Equal}}$ = 26, $n_{\text{Control}}$ = 19) and 355 children ($n_{\text{Think Equal}}$ = 237, $n_{\text{Control}}$ = 118) in total participated in this study within the Australian States of Victoria and Queensland.

Teachers in both Think Equal and Control conditions were asked at the beginning (i.e., T1) and end (i.e., T2) of the academic year (April–December 2019) to report on children's demographics characteristics (i.e., age, gender, English-language proficiency) and children's social and emotional behaviour and learning using widely-used, validated assessments. Subscales included Emotion Regulation and Emotion Liability or "Emotion Dysregulation" from the Emotion Regulation Checklist[2], Anger/Aggression, Anxious/Withdrawn, and Social

---

[1] Sullivan, G. M. (2011). Getting off the "gold standard": randomized controlled trials and education research. *Journal of Graduate Medical Education, 3*, 285–289. https://doi.org/10.4300/JGME-D-11-00147.1

[2] Shields, A., & Cicchetti, D. (1997). Emotion regulation among school-age children: the development and validation of a new criterion Q-sort scale. *Developmental Psychology, 33*, 906–916.

Competence from the Social Competence and Behavior Evaluation-30[3], and Effortful Control, Surgency or "Extraversion", and Negative Affect from the Early Childhood Behavior Questionnaire[4]. These subscales were reliable at each timepoint and strongly associated between timepoint (see Table 1).

**Measures**
The definitions for each measure used in this study e.g., emotion regulation, effortful control etc, are listed in Appendix 1, while the subscales operationalized by the items in each measure are stated below as:

**Emotion Regulation Checklist**
- Emotion Regulation Shields, A., & Cicchetti, D. (1997). Emotion regulation among school-age children: the development and validation of a new criterion Q-sort scale. *Developmental Psychology, 33*, 906–916.
- 24 items total
- Likert-type scale: "Never," "Sometimes," "Often," "Almost Always"
- Emotion Regulation (8 items)
- Emotion Liability (i.e., Emotion Dysregulation; 16 items)

**Social Competence and Behavior Evaluation-30**
- LaFreniere, P. J., & Dumas, J. E. (1996). Social competence and behavior evaluation in children ages 3 to 6 years: The short form (SCBE-30). *Psychological Assessment, 8*, 369–377. https://doi.org/10.1037/1040-3590.8.4.369
- 30 items total
- Likert-type scale: "Never," "Rarely," "Sometimes," "Often," "Frequent," "Always"
- Anger/Aggression (10 items)
- Anxious/Withdrawn (10 items)
- Cooperation/Sensitive (i.e., Social Competence; 10 items)

**Early Childhood Behavior Questionnaire-Very Short Teacher Form**
- Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: the early childhood behavior questionnaire. *Infant Behavior & Development, 29*, 386–401. https://doi.org/10.1016/j.infbeh.2006.01.004
- 24 items total
- "Never," "Very rarely," "Less than half the time," "About half the time," "More than half the time," "Almost always," "Always"
- Effortful Control (12 items)
- Surgency (i.e., Extraversion; 12 items)
- Negative Affect (12 items)

---

[3]LaFreniere, P. J., & Dumas, J. E. (1996). Social competence and behavior evaluation in children ages 3 to 6 years: The short form (SCBE-30). *Psychological Assessment, 8*, 369–377. https://doi.org/10.1037/1040-3590.8.4.369

[4] Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: the early childhood behavior questionnaire. *Infant Behavior & Development, 29*, 386–401. https://doi.org/10.1016/j.infbeh.2006.01.004

## Analysis

The breakdown of gender in this study included 59% boys and 41% girls ranging in age from 3 – 5 years with an average age of 4.45 years ($SD$ = 0.72). English proficiency in the children was at an overall rate was 92% ranging from 67 to 100%. Age differences between the two groups at T1 were explored and revealed that children in the Think Equal group were younger, more likely to be boys, and were reported as angrier and more aggressive, emotionally dysregulated, and extraverted by 0.27 to 0.40 $SD$ units (see Table 2). To address baseline differences and create a pseudo counterfactual to promote causal inference, Inverse Probability of Treatment Weighting (IPTW) was used, which uses Propensity scoring to re-weight coefficients within analyses. Covariate adjustment was also used. Table 3 shows the comparison of both groups at T1 after using IPTW revealing minimal differences between the groups at T1.

## Caveats

By T2, 35 centres ($n_{\text{Think Equal}}$ = 23, $n_{\text{Control}}$ = 11) and 255 children ($n_{\text{Think Equal}}$ = 179, $n_{\text{Control}}$ = 76) remained in the study. Some of the centres (i.e., 3 Think Equal and 8 control) that started the study dropped out for an overall attrition rate of 24% and a differential attrition rate of 31%. Overall, the effect of attrition appeared to be minimal based on comparing T1 scores and demographic characteristics among those who remained at T2 and those who left. All subsequent results refer to those students that remained in the study.

## Results

Differences between Think Equal and Control Groups at T2 after controlling for T1, age, English language proficiency, and gender were explored using multi-level modelling nesting children within classroom and classroom within school. Intraclass correlation coefficients revealed minimal to substantial variation in children's T2 scores due to classroom and school membership, justifying our choice in modeling.

As shown in Figure 1, the impact of Think Equal depended on children's scores at the beginning of the year for emotion regulation, emotion dysregulation, anxious and withdrawn, extraversion, and negative affect and did not depend on beginning of the year scores for effortful control. No statistically significant differences between the Think Equal and control group were found for children's angry and aggressive behaviors ($d$ = 0.21) or their social competence ($d$ = 0.35) despite sizeable differences between the groups.

For the five measures that depended on the children's initial scores, the black vertical line on the graphs indicates the cut-off initial score at which the difference first becomes significant between the Think Equal and control groups.

Figure 2 further explores the results from Figure 1. For each measure, the mean scores of those children above and below the T1 cut-off scores for both the Think Equal and control groups are shown. The percentage of children in each range of initial scores above and below the cut-off score is also shown to identify the proportion of children that benefitted from the Think Equal program the most. For example, 55% of children scored below the cut-off score of 26.2 for emotion regulation and children with these lower scores benefitted from the Think Equal program. For those measures in which the initial score did not affect the results, a comparison of the means in the Think Equal and control groups for all children is shown. That is, there is no need to split the results based on cut-off scores for these measures.

## Conclusion

Overall, statistical analysis demonstrates that the children that participated in the Think Equal program were more emotionally regulated ($d$ = 0.38), less emotionally dysregulated ($d$ = 0.32), less anxious and withdrawn ($d$ = 0.33), demonstrated greater effortful control ($d$ = 0.45), demonstrated greater extraversion ($d$ = 0.24), and had lower negative affect ($d$ = 0.37). These findings provide very strong evidence-based validation for the extensive benefits of the Think Equal program within the Australian context.

## Additional information

- Dr Susan Emmett - s.emmett@federation.edu.au
- Dr Lynne Reeder -  l.reeder@federation.edu.au

**Table 1**

Internal Consistency Among and Associations Between Eight T1 and T2 Outcome Variables

| Measure | Items | $\alpha_{T1}$ | $\alpha_{T2}$ | r |
|---|---|---|---|---|
| Emotion Regulation | 8 | .86 | .84 | .55 |
| Emotion Dysregulation | 16 | .89 | .89 | .72 |
| Anger and Aggressive | 10 | .88 | .93 | .66 |
| Anxious and Withdrawn | 10 | .91 | .88 | .53 |
| Social Competence | 10 | .92 | .91 | .56 |
| Effortful Control | 12 | .80 | .81 | .48 |
| Extraversion | 12 | .85 | .78 | .65 |
| Negative Affect | 12 | .76 | .71 | .49 |

**Table 2**

Comparison of all Measures for the Think Equal and Control Groups at Baseline T1 prior to Inverse Probability of Treatment Weighting.

| | Control ($n_1$ = 76) | Think Equal ($n$ = 179) | |
|---|---|---|---|
| | % | % | Cox's Index |
| Gender[a] | 47% | 59% | 0.30 |
| English[b] | 92% | 92% | 0.03 |
| | M (SD) | M (SD) | d |
| Age[c] | 4.71 (0.54) | 4.33 (0.68) | 0.59 |
| Emotion Regulation | 25.04 (5.22) | 25.04 (4.00) | 0.00 |
| Emotion Dysregulation | 25.57 (7.53) | 27.62 (7.45) | 0.27 |
| Anger and Aggressive | 17.91 (9.00) | 21.05 (7.27) | 0.40 |
| Anxious and Withdrawn | 22.84 (10.20) | 23.01 (7.95) | 0.02 |
| Social Competence | 37.42 (11.03) | 37.97 (8.35) | 0.06 |
| Effortful Control | 5.38 (0.95) | 5.25 (0.88) | 0.14 |
| Extraversion | 4.06 (1.22) | 4.39 (1.02) | 0.30 |
| Negative Affect | 2.24 (1.13) | 2.29 (0.79) | 0.06 |

*Notes.* Effect sizes were Cohen's *d* for continuous variables and Cox's Index for categorical variables; [a] Male = 1. [b] % "good". [c] age in years.

**Table 3**
Comparison of all Measures for the Think Equal and Control Groups at Baseline T1 after using Inverse Probability of Treatment Weighting.

| | Control ($n_l$ = 76) | Think Equal ($n$ = 179) | |
|---|---|---|---|
| | % | % | Cox's Index |
| Gender[a] | 56% | 56% | 0.00 |
| English[b] | 92% | 91% | 0.07 |
| | M (SD) | M (SD) | d |
| Age[c] | 4.34 (0.67) | 4.71 (0.54) | 0.03 |
| Emotion Regulation | 26.80 (7.43) | 27.20 (7.69) | 0.05 |
| Emotion Dysregulation | 19.74 (7.78) | 20.66 (8.05) | 0.12 |
| Anger and Aggressive | 21.99 (8.18) | 22.46 (8.47) | 0.06 |
| Anxious and Withdrawn | 38.40 (8.85) | 38.20 (9.15) | 0.02 |
| Social Competence | 5.27 (0.87) | 5.30 (0.91) | 0.03 |
| Effortful Control | 4.43 (1.05) | 4.37 (1.10) | 0.06 |
| Extraversion | 2.16 (0.86) | 2.27 (0.90) | 0.12 |
| Negative Affect | 25.52 (4.27) | 25.01 (4.42) | 0.12 |

*Notes.* Effect sizes were Cohen's *d* for continuous variables and Cox's Index for categorical variables; [a] Male = 1. [b] % "good". [c] age in years.

**Figure 1** Interaction plots for significant measures showing the difference between treatment and control groups depending on the initial score. The black vertical line for each measure shows the point at which the Think Equal group outperformed the Control Group.
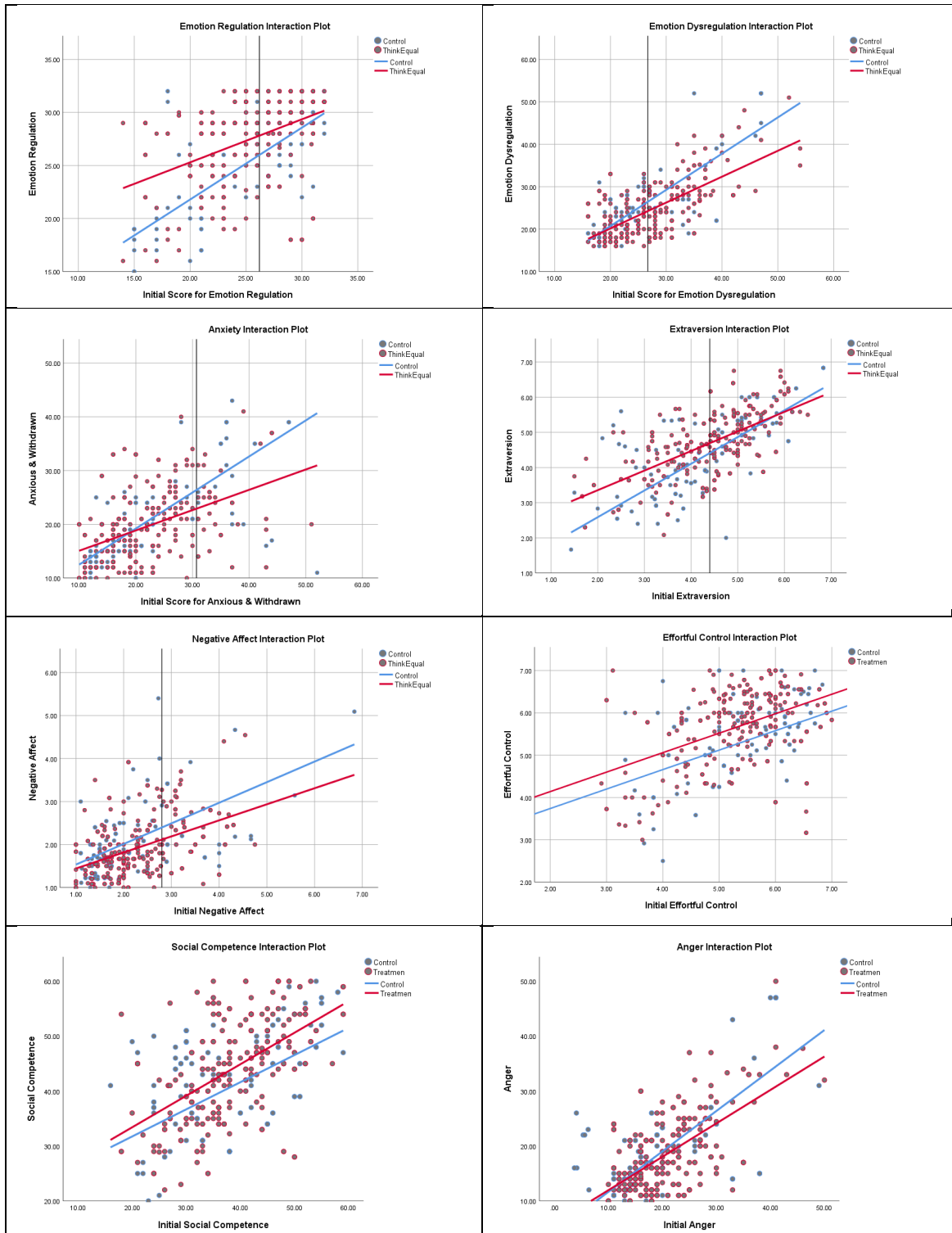
**Figure 2** Bar charts for all measures showing the difference between treatment and control groups.
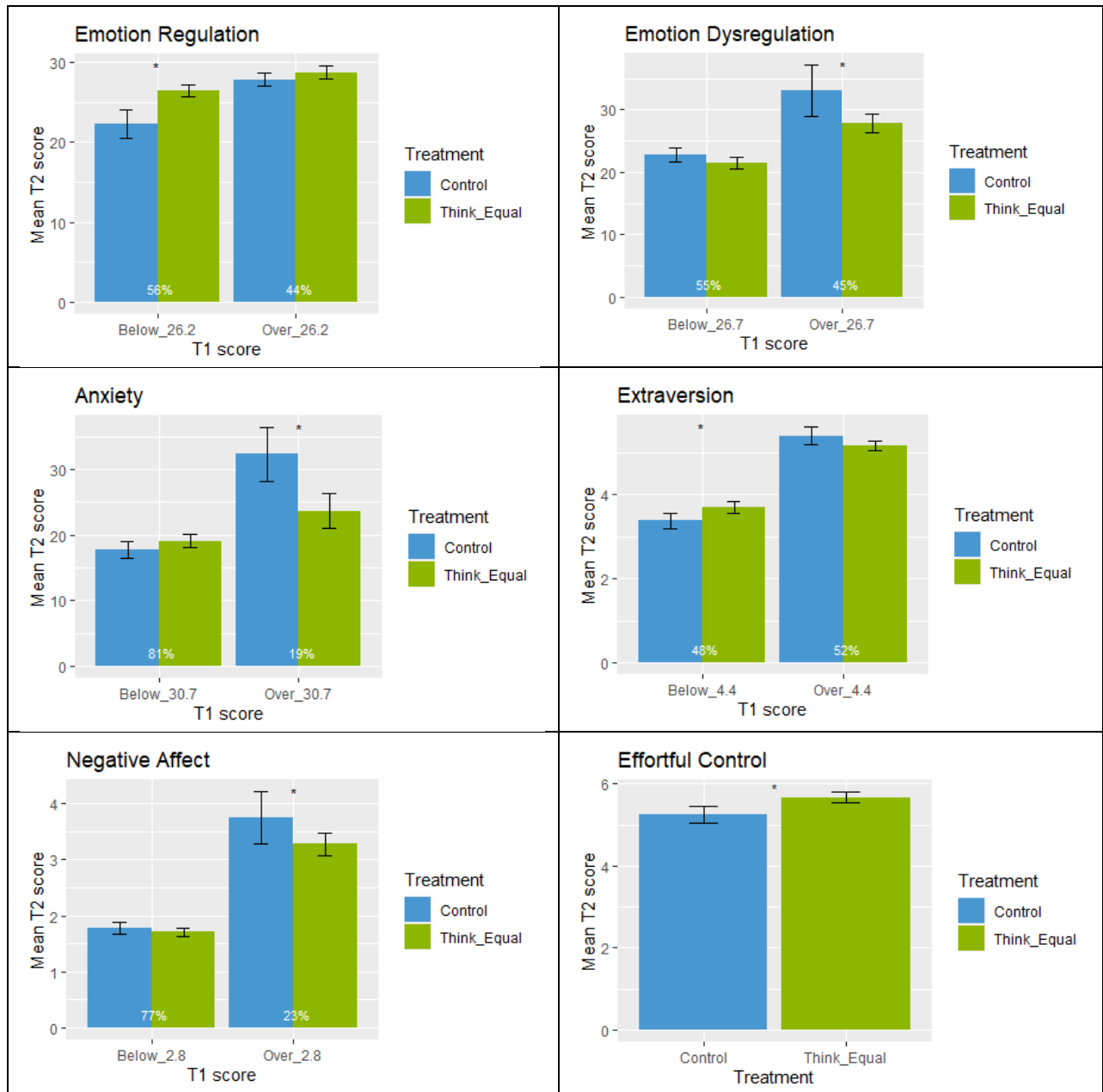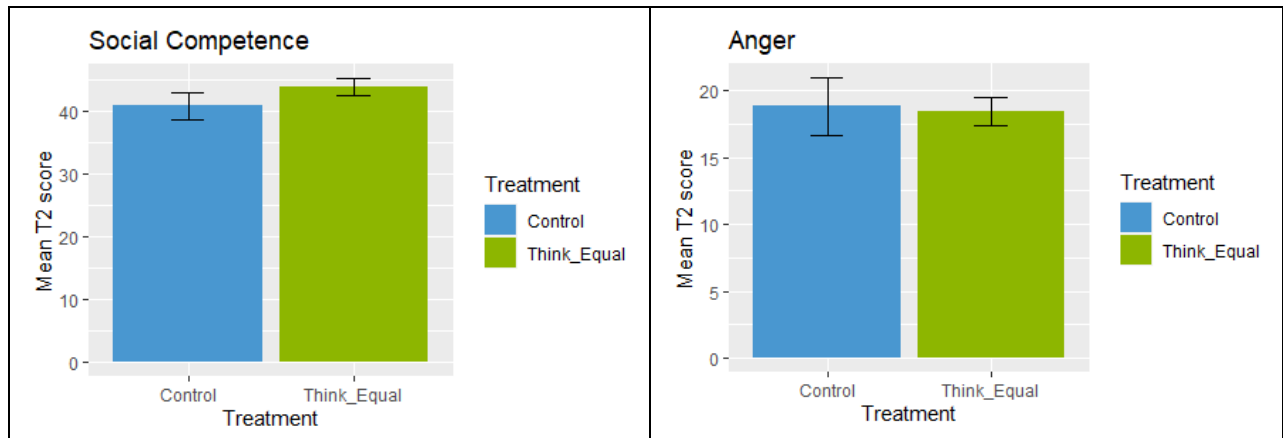
**Note:** The error bars refer to +/- 2 std errors from the mean. The '*' refers to the initial scores identified in Figure 1 in which the treatment outperforms the control group. The values in white at the bottom of the bars refers to the percentage of children in that range of initial scores.

**Appendix 1**

**Definitions for each measure comprise:**

- **Emotion regulation** involves the cognitive and behavioral processes children use in changing the emotions they have, when they have them, and how they experience and express them (Gross, 1998). A child who is emotionally regulated, as is the case for our paper, is one whose emotions are regulated in such a way to effectively achieve their goals or social goals.

- **Emotion liability** describe rapid changes in emotions often accompanied with high intensity. Emotion regulation describes how emotions change, whereas emotion liability describes how emotions are expressed. A child who scores high in emotion liability is often dysregulated as they struggle to meet their goals or social goals. A child who scores low in liability is also emotionally regulated, which is why the two constructs are correlated yet still distinct.

- **Anger/Aggression, Cooperation/Sensitive, Anxious/Withdrawn** are specific to the SCBE-30. The authors note that Cooperation/Sensitive describes prosocial behaviors and is sometimes just refereed to as Social Competence, Anger/Aggression as externalizing and Anxious/Withdrawn as internalizing to align with other theories. We see overlap between SCBE subscales and ERC subscales, which makes sense given that the SCBE also captures emotion regulation and expression. Keep in mind that these are a classification. All children have degrees of each.

- **Effortful control** is similar conceptually to emotion regulation in playing a regulatory functioning but broader in that it describes a child who purposefully deploys their attentional network to inhibit behavioral responses (Rothbart, Ahadi, & Hershey, 1994). Effortful control is sometimes called executive control to align it with aspects of executive functions (i.e., attention, inhibitory control, and working memory), but effortful control and the other dimensions of temperament are older terms. I'd argue that effortful control (and broadly, executive functioning) is part of what's happening under the hood when a child is regulating their emotion. But effortful control extends to other aspects of functioning as well.

- **Negative affect** is straightforward and is really describing patterns of high and frequent fear, anger, sadness, and discomfort and similar to Anger/Aggression and Anxious/Withdrawn. It can be thought of as an action tendency, so a child who is likely to respond fearfully, angerly, etc.

- **Surgency** is similar to how people describe extraversion: positive emotionality, high activity level, impulsivity, and risk-taking. Sometimes, people will describe it as positive affect and approach behaviors, especially when describing babies. Surgency is often used when discussing temperament with young children, whereas extraversion is often used when discussing personality with older children and adults. Effortful control, negative affect, and surgency are altogether part of Rothbart's model of temperament