
# Analysis of an In-School Mental Health Services Model for K-12 Students Requiring Intensive Clinical Support

## A White Paper Report on Tier 3 School-Based Mental Health Programming

Amanda M. Dettmer, PhD
Yale Child Study Center
September 18, 2023

## Abstract

Emotional, behavioral, and mental health challenges make it difficult for many children and adolescents to engage and succeed at school. Research indicates that at least 20% of all children and adolescents have been diagnosed with one more mental health disorders. Behavioral problems, anxiety, and depression are the most diagnosed mental health issues, and they often co-occur. Moreover, these conditions are being diagnosed at increasingly younger ages. In the past several years there has been a rise in the number of adolescents and young adults with serious mental health issues such as major depression and suicidal ideation, and the COVID-19 pandemic has exacerbated mental health problems for children and adolescents. Schools offer promise for providing intensive clinical support to the most at-risk students, and schools are necessary environment to explore the implementation of multi-modal youth mental health services. This paper provides an analysis of an intensive, in-school mental health services model developed and implemented by Effective School Solutions (ESS), a New Jersey based provider of high acuity school based mental health services for K-12 students. We analyze this multi-modal model for its effectiveness in improving educational outcomes for over 3,000 students identified as requiring intensive clinical mental health support across the 2021-22 school year. This analysis reveals that those students receiving High- versus Low-fidelity programming (i.e., multiple sessions per week for at least half of the school year versus for less than half of the school year) had better educational outcomes. Students receiving High-fidelity programming had greater improvements in grade point average (GPA) and greater reductions in absences across the school year. A higher number of in-school clinical sessions per week significantly predicted a greater increase in GPA and a greater reduction in total disciplinary incidents (including out of school suspensions) across the school year. This report provides initial promising evidence that in-school intensive mental health clinical services yield positive effects on students' educational outcomes. Though future research is needed to validate and extend these findings, schools may consider implementing such services onsite to meet students where they are and to optimize students' mental, behavioral, and educational well-being.

**Keywords** mental health, in-school services, clinical care, K-12

# Introduction

Emotional, behavioral, and mental health challenges make it difficult for many children and adolescents to engage and succeed at school. Research indicates that nearly 50% of adolescents have had a mental disorder at some point in their lives, with over 20% experiencing severe impairment and/or distress (National Institute of Mental Health, 2021). Behavioral problems, anxiety, and depression are the most commonly diagnosed mental health issues and they often co-occur; these conditions are being diagnosed at increasingly younger ages (CDC, 2019). Of great concern, in the past several years there has been a rise in the number of adolescents and young adults with serious mental health issues such as major depression and suicidal ideation (Twenge et al., 2019). Moreover, the COVID-19 pandemic has exacerbated mental health problems for many children and adolescents (Beal, 2021; Ng et al., 2022).

The efficacy of a multi-modal approach to intensive clinical mental health care is supported by extensive research (Diamond & Josephson, 2005; Matta, 2014; Rones & Hoagwood, 2000; Zhou et al., 2015, 2018). However, access to intensive clinical support for mental health needs is constrained by many factors, including availability and cost of services and caregivers' and students' available time outside of work and school to attend appointments, among many others. Providing intensive, clinical mental health support for the most at-risk students in schools has the potential to mitigate many of these constraints by meeting students where they are most of the day – at school. When provided as part of school district services, equity to mental health services is enhanced. Schools are an appropriate and necessary environment to explore the implementation of youth mental health services (Matta, 2014; Shechtman, 2007), but few studies have examined the educational impacts of in-school services.

This paper provides an analysis of an intensive, in-school mental health services model developed by Effective School Solutions (ESS) for K-12 students. ESS partners with school districts across the U.S. to provide mental and behavioral health support services in schools. Services are provided by licensed clinicians using a multi-tiered system of support (MTSS) framework. One tier, Tier 3, provides services to the most at-risk students: those identified via screening by mental health professionals as needing intensive, in-school clinical support. This analysis focuses on educational outcomes for students receiving ESS' Tier 3 services. By comparing students receiving variable amounts and types of services, the impact of program fidelity is assessed and described.

# Methods

## Participants

This analysis included all Tier 3 students aged 9yrs and older from school districts in 9 different states who received ESS services (N=3,119 total students). All data were provided to the author in anonymized fashion by ESS; no personally identifiable information was included.

## ESS Tier 3 In-School Mental Health Services Model

The ESS Tier 3 program is provided to students referred by school district staff members to receive intensive, in-school clinical mental health support. While referrals generally originated with existing school district referral processes, the placement into Tier 3 services was validated by a full bio/psycho/social clinical assessment administered by the provider. Clinician-provided services included Family Sessions, Learning Strategies Sessions, Parent Support Groups, Group Services, Individual Sessions, School Meetings, Structured Lunch, Therapeutic

Check-Ins, *pro re nata* (PRN) Sessions, Classroom Observation/Interventions, Home Visits, Classroom Psychoeducation, and Other sessions. Students were classified as having received *High-fidelity programming* if they received at least 3 non-administrative services per week plus 1 family session in the past 2 weeks for at least 50% of the school year (n=661), and as having received *Low-fidelity programming* if they received this same combination of services for less than 50% of the school year (n=2,458).

## Primary Outcomes

The following educational outcomes were assessed monthly as a function of delivery of Tier 3 services:

1. GPA (improvements from baseline [Y/N]; absolute GPA during marking periods; and changes in GPA [percent change from baseline]);
2. Absences (improvements from baseline [Y/N]; absolute number of absences during marking periods; and changes in absences [percent change from baseline]);
3. Disciplinary Incidents (improvements from baseline [Y/N]; absolute number of individual and total disciplinary incidents during marking periods; and changes in individual and total disciplinary incidents [percent change from baseline]);
4. Out of district placements (ODP) (high risk [Y/N]);
5. Total number of major incidents across the school year (including aggression, elopement, child protective services, self-harm, danger to self, and others);
6. Student Wellness Ratings (average ratings at the beginning and the end of the year; and changes in ratings [percent change from baseline])

## Analysis

Initial descriptive chi-square analyses were conducted to probe for associations between fidelity (high and low) and demographic characteristics (sex, gender identity, and racial identity). Because only racial identity was associated with program fidelity ($\chi^2$=13.9, p=0.003), racial identity was controlled for in all subsequent analyses. Next, three-way crosstab and chi-square analyses were conducted to test for associations between high/low program fidelity and dichotomous outcomes (Y/N improvements in GPA, absences, disciplinary incidents between baseline and marking periods [MPs]; risk of ODP; and positive/neutral or negative wellness ratings) while controlling for racial identity. Following these, multivariate analyses of covariance (MANCOVAs) were used to test for program fidelity differences in scaled outcomes after controlling for racial identity. Here, MANCOVAs were run in four clusters owing to non-uniformity of missing data across outcomes: 1) GPA (baseline, MP, and % change), 2) absences (baseline, MP, and % change), 3) disciplinary incidents (baseline, MP, and % change), and 4) wellness ratings (baseline, MP, and % change). A univariate analysis of covariance (ANCOVA) was used to test for program fidelity differences in total major incidents, which was only recorded (i.e., summed) once over the entire school year. To probe for a potential "dosing effect," after controlling for racial identity regression analyses examined the extent to which total average sessions per week (combination of group sessions, learning strategies sessions, and family sessions) predicted changes in GPA, absences, disciplinary incidents, and wellness ratings across the school year, as well as total major incidents across the year, using the F-test for the $\Delta R^2$ value. For the regressions, the additional predictive values of sex and gender identity were tested. For all disciplinary outcomes, data were filtered to include only those students who had >0 incidents at the start of the school year (baseline). For the above analyses, statistical significance was set at p=0.05. Finally, a follow-up sensitivity analysis was conducted whereby, for High- and Low-fidelity groups separately, partial correlations were used to examine the relation between total average sessions per week and outcomes of interest while controlling for

racial identity. For the partial correlations, statistical significance was set at p=0.003 to adjust for multiple comparisons.

# Results

## Demographics

The number of High- and Low-fidelity students categorized by sex, gender identity, and racial identity is given in Table 1, along with the percentages of the total sample in parentheses. The results of the chi-square analyses examining associations between demographic factors and program fidelity are also indicated. Only racial identity was significantly associated with program fidelity ($\chi^2$=13.9, p=0.003, n=3,044; Figure 1; Table 1), though the effect size was very small ($\phi$=0.07).

*Table 1. ESS Tier 3 Program Fidelity (High vs. Low) associations with student demographics.*

| Program Fidelity | | Sex[a] | | |
|---|---|---|---|---|
| | Female | Male | Missing/Other | |
| High | 300 (9.6%) | 359 (11.5%) | 2 (0.1%) | |
| Low | 1098 (35.3%) | 1354 (43.4%) | 6 (0.2%) | |
| Total | 1398 (44.8%) | 1713 (54.9%) | 8 (0.3%) | |
| | | Gender Identity | | |
| | Cisgender Female | Cisgender Male | Transgender/Non-Binary/ Gender Fluid/Other | |
| High | 245 (7.9%) | 368 (11.8%) | 48 (1.5%) | |
| Low | 926 (29.7%) | 1313 (42.1%) | 219 (7.0%) | |
| Total | 1171 (37.5%) | 1681 )53.9%) | 267 (8.6%) | |
| | | Racial Identity[b,c] | | |
| | African American | Hispanic/Latino | AAPI/AIAN/Multi-Racial/Other | White |
| High | 104 (3.4%) | 128 (4.2%) | 94 (3.1%) | 282 (10.4%) |
| Low | 525 (17.2%) | 520 (17.1%) | 335 (11.0%) | 1022 (33.6%) |
| Total | 629 (20.7%) | 648 (21.3%) | 429 (14.1%) | 1338 (44%) |

[a]Sex assigned at birth; differs from gender identity

[b]AAPI = Asian American/Pacific Islander; AIAN = American Indian/Alaska Native

[c]Significant association between racial identity and program fidelity ($\chi^2$=13.9, p=0.003)

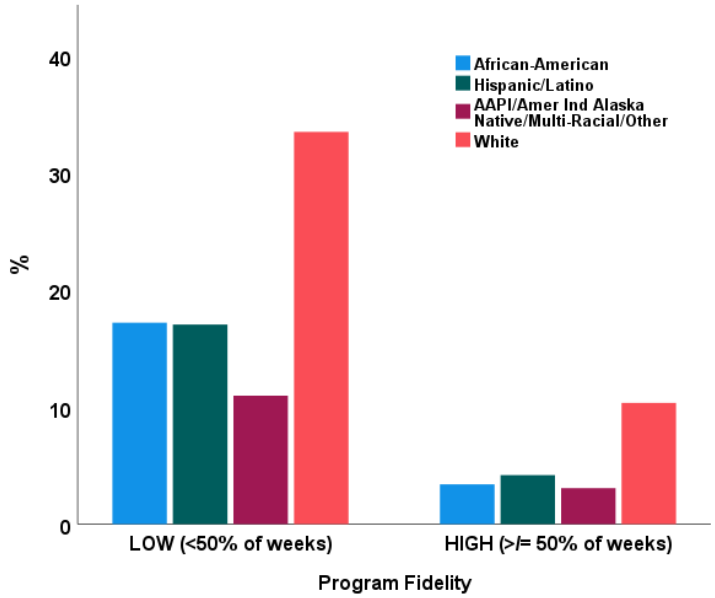


*Figure 1. Significant association between program fidelity and racial identity.*

## Associations between program fidelity and outcomes of interest (dichotomous data)

Significant associations between program fidelity and GPA improvement ($\chi^2$=12.9, p<0.001, ϕ=0.08, n=2,088), and program fidelity and high risk of ODP ($\chi^2$=6.1, p=0.13, ϕ=0.07, n=1,404) were observed, though effect sizes were very small. A greater proportion of High-fidelity students had improvements in GPA compared with Low-fidelity students (Figure 2), but a greater proportion of Low fidelity students were at high risk of ODP (Figure 3). No other associations were observed.

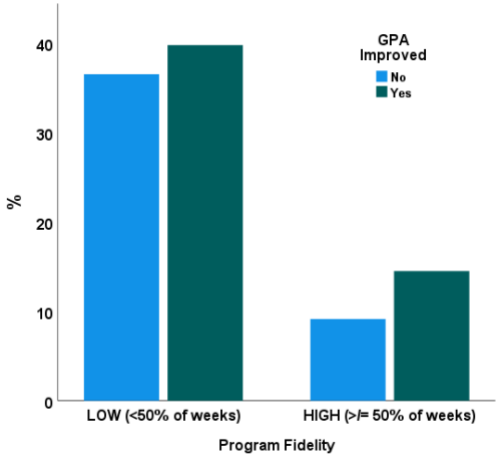


*Figure 2. Significant association between program fidelity and GPA improvement ($\chi^2$=12.9, p<0.001).*

## Program fidelity group differences in outcomes of interest (scaled data)

Program fidelity had a significant impact on many outcomes of interest (Table 2; Figure 4). High-fidelity students had a higher marking period (MP) GPA, fewer baseline and MP absences, fewer out of school suspensions (OSS) at baseline and across MPs, and higher self-ratings of wellness averaged across all weeks of the school year. High-fidelity students also had a greater % change in GPA and a smaller % change in absences across the school year. No interactions with sex, gender identity, or racial identity were observed.

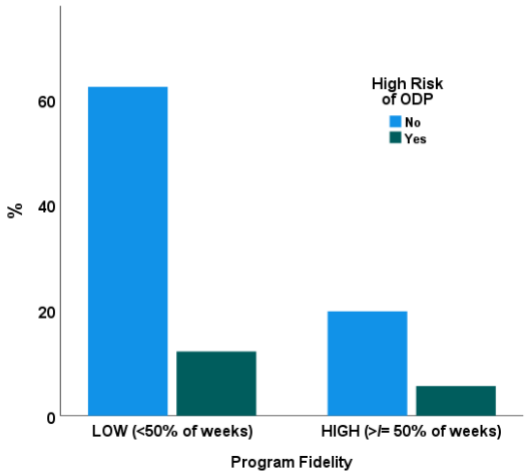


*Figure 3. Significant association between program fidelity and ODP risk ($\chi^2$=6.1, p<0.013).*

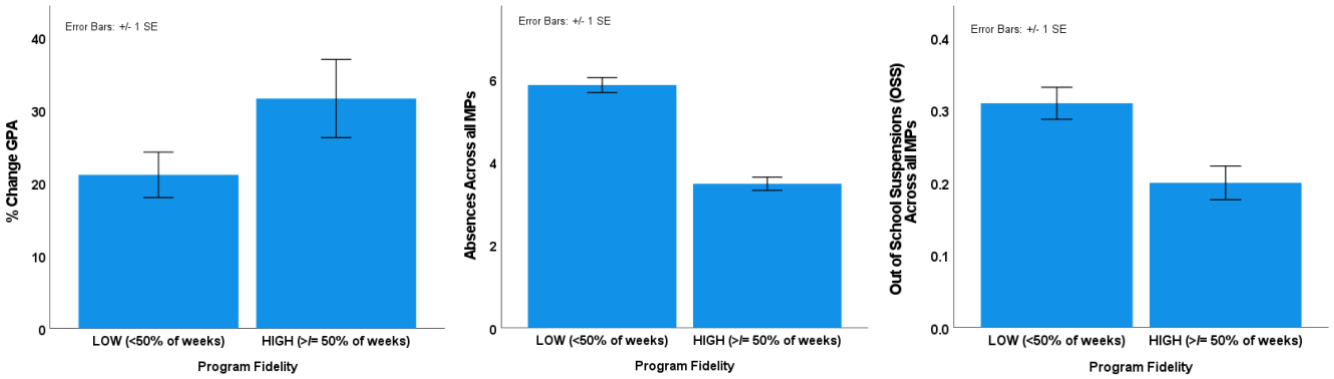


*Figure 4. High-fidelity students had greater average increases in GPA (left) and smaller average increases in absences (center) across the school year, and fewer out of school suspensions (OSS, right) across all marking periods (MPs). (All p's <0.05; see also Table 2.)*

*Table 2. Impacts of program fidelity on educational outcomes.*

| Outcome | n | t | p | Effect size \|d\| | Direction of effect[a] |
|---|---|---|---|---|---|
| Baseline GPA | 2,108 | -0.96 | 0.34 | N/A | N/A |
| **MP GPA** | **2,348** | **-5.23** | **<0.001** | **0.25** | **H > L (2.6±0.05 vs. 2.3±0.03)** |
| **% Change GPA[b]** | **1,815** | **-1.64** | **0.05** | **0.09** | **H > L (31.6±5.4% increase vs. 21.1±3.1% increase)** |
| **Baseline Absences** | **2,211** | **4.39** | **<0.001** | **0.19** | **L > H (6.7±0.2 vs. 4.8±0.3)** |
| **MP Absences** | **2,298** | **9.86** | **<0.001** | **0.35** | **L > H (5.7±0.2 vs. 3.5±0.2)** |
| **% Change Absences[b]** | **1,669** | **1.89** | **0.03** | **0.11** | **L > H (41.3±5.7% increase vs. 19.4±8.0% increase)** |
| Baseline OSS | 371 | 1.50 | 0.13 | N/A | N/A |
| **MP OSS** | **366** | **2.41** | **0.02** | **0.21** | **L > H (0.9±0.1 vs. 0.6±0.08)** |
| Baseline Total Disciplinary Events[c] | 371 | 0.34 | 0.74 | N/A | N/A |
| MP Total Disciplinary Events[c] | 363 | 0.67 | 0.50 | N/A | N/A |
| % Change Total Disciplinary Events[c] | 363 | 0.60 | 0.55 | N/A | N/A |
| Wellness Rating[d] (start of year) | 2,192 | 0.38 | 0.70 | N/A | N/A |
| Wellness Rating (end of year) | 730 | 0.04 | 0.97 | N/A | N/A |
| **Average Wellness Rating (entire year)** | **2,987** | **2.56** | **0.01** | **0.12** | **L > H (2.54±0.01 vs. 2.48±0.02)** |
| % Change Wellness Rating | 680 | 1.08 | 0.28 | N/A | N/A |
| **Total Major Incidents** | **3,119** | **-5.93** | **<0.001** | **0.30** | **H > L (1.1±0.08 vs. 0.6±0.03)** |

[a]H=High-Fidelity, L=Low-fidelity; [b]One-tailed t-tests were performed for % change scores based on the results of the two-tailed MANCOVAs for group differences in baseline and/or MP scores; [c]For disciplinary outcomes, only students with baseline disciplinary events >0 were included; [d]For all wellness rating measures, a lower score indicates better wellness

## Outcomes predicted by session delivery

After controlling for racial identity, average total weekly session delivery significantly predicted multiple outcomes of interest (Table 3), though effect sizes were very small. More weekly services (group + family sessions, learning strategies) predicted an increase in GPA, a decrease in absences, and a reduction in total disciplinary events, but a marginally lower average self-reported wellness rating across the school year and marginally more reported major incidents. For all significant outcomes, the effects were very small (all $R^2$ values <0.02; Table 3).

After controlling for racial identity, the sensitivity analysis showed significant correlations between average weekly session delivery and several educational outcomes for Low-fidelity students only. Using the stringent p=0.003 criterion, more weekly therapeutic sessions were significantly correlated with MP GPA (r=0.18, p<0.001), MP absences (r= -0.13, p<0.001), MP out of school suspensions (OSS; r= -0.12, p<0.001), and total major incidents across the school year (r=0.10, p<0.001). These correlations were significant for Low-fidelity students only; no significant correlations were observed for High-fidelity students.

## Strengths and Limitations

This study is strengthened by the large sample size and diversity in the population in terms of gender identity, racial identity, and geographic distribution. A limitation of this study is that it was not a randomized control trial, so there is no control group; thus, these early findings should be interpreted with caution. Follow up studies are needed to replicate and extend these findings.

*Table 3. Educational outcomes predicted by session delivery.*

| Outcome | β | $R^2$ | p | Significant Interactions[b] |
|---|---|---|---|---|
| **Baseline GPA** | **0.03** | **0.002** | **0.045** | **Racial identity ($\beta$=0.13, $\Delta R^2$=0.02, p<0.001)** |
| **MP GPA** | **0.13** | **0.03** | **<0.001** | **Racial identity ($\beta$=0.16, $\Delta R^2$=0.02, p<0.001)** |
| **% Change GPA** | **4.6** | **0.004** | **0.005** | **Sex ($\beta$=-13.9, $\Delta R^2$=0.004, p=0.01)** |
| **Baseline Absences** | **-0.5** | **0.006** | **<0.001** | **Sex ($\beta$=-1.0, $\Delta R^2$=0.002, p=0.02)** |
| **MP Absences** | **-0.7** | **0.03** | **<0.001** | **Sex ($\beta$=-0.9, $\Delta R^2$=0.004, p=0.002)** |
| **% Change Absences** | **-5.9** | **0.002** | **0.045** | N/A |
| Baseline OSS | -0.07 | 0.006 | <0.001 | **Sex ($\beta$=0.16, $\Delta R^2$=0.003, p=0.009) Racial identity ($\beta$=-0.11, $\Delta R^2$=0.01, p<0.001)** |
| **MP OSS** | **-0.06** | **0.02** | **<0.001** | **Sex ($\beta$=0.16, $\Delta R^2$=0.01, p<0.001) Racial identity ($\beta$=-0.06, $\Delta R^2$=0.01, p<0.001)** |
| Baseline Total Disciplinary Events[c] | -0.10 | 0.01 | <0.001 | **Sex ($\beta$=0.27, $\Delta R^2$=0.005, p=0.002) Racial identity ($\beta$=-0.20, $\Delta R^2$=0.02, p<0.001)** |
| MP Total Disciplinary Events[c] | -0.06 | 0.01 | <0.001 | **Sex ($\beta$=0.20, $\Delta R^2$=0.005, p<0.001) Racial identity ($\beta$=-0.10, $\Delta R^2$=0.01, p<0.001)** |
| **% Change Total Disciplinary Events[c]** | **-9.8** | **0.016** | **0.017** | N/A |
| Wellness Rating[d] (start of year) | -0.01 | <0.001 | 0.59 | N/A |
| Wellness Rating (end of year) | 0.003 | <0.001 | 0.80 | N/A |
| **Average Wellness Rating (entire year)** | **-0.02** | **0.003** | **0.002** | **Sex ($\beta$=-0.11, $\Delta R^2$=0.01, p<0.001) Gender identity ($\beta$=0.04, $\Delta R^2$=0.002, p<0.03)** |
| % Change Wellness Rating | 0.9 | 0.001 | 0.36 | |
| **Total Major Incidents** | **0.1** | **0.02** | **<0.001** | **Racial identity ($\beta$=-0.11, $\Delta R^2$=0.006, p<0.001)** |

[a]Average total weekly session delivery = total number of group + family sessions, learning strategies each week

[b]Sex: Female=1, Male=2; Gender Identity: Cisgender Female=1, Cisgender Male=2, Transgender/Non-Binary/Gender Fluid/Other=3; Racial Identity: African American=1, Hispanic/Latino=2, AAPI/American Indian Alaska Native/Multi-Racial/Other=3, White=4

[c]For disciplinary outcomes, only students with baseline disciplinary events >0 were included

[d]For all wellness rating measures, a lower score indicates better wellness

# Conclusions

This analysis revealed multiple positive effects on educational outcomes for students receiving High-fidelity ESS Tier 3 program implementation. High-fidelity program implementation resulted in better academic and non-academic outcomes compared to Low-fidelity program implementation. Students receiving High-fidelity programming fared better in terms of GPA improvement, reductions in absences, fewer disciplinary incidents (especially out of school suspensions), and out of district placement (ODP) risk across the 2021-22 school year. Moreover, a higher number of average sessions per week significantly predicted positive outcomes, and results indicate that there is a dose-response for students receiving Low-fidelity programming, such that those on the higher end of the spectrum (i.e., at least 3 non-administrative sessions per week and 1 family session in the previous 2 weeks closer to 50% of weeks across the school year) see greater benefits, but for those receiving these services at least 50% of the weeks across the school year, more is not necessarily better. These findings have important implications for policymakers, specifically school districts and boards of education, with respect to implementation of much-needed mental health services in schools. Further research is needed to validate and extend these findings, including studies incorporating control groups in either randomized control trials or quasi-experimental designs. Collectively, there is initial promising evidence for ESS Tier 3 High-fidelity program implementation on educational outcomes across a diverse sample of K-12 students.

# References


Beal, J. A. (2021). Impact of the covid-19 pandemic on mental health of children, adolescents, and young adults. *MCN: The American Journal of Maternal/Child Nursing*, *46*(4), 237.

Diamond, G & Josephson, A. (2005). Family-based treatment research: A 10-year update. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*(9), 872-887.

Matta, A. R. (2014). Efficacy of school-based group therapy with children and adolescents: A meta-analytic review. Dissertation Abstracts International: Section B: The Sciences and Engineering. ProQuest Information & Learning. Retrieved from http://search.ebscohost.com.ezp-prod1.hul.harvard.edu/login.aspx?direct=true&db=psyh&AN=2015-99020-510&site=ehost-live&scope=site

National Institute of Mental Health. (2021). *Mental illness.* Retrieved from https://www.nimh.nih.gov/health/statistics/mental-illness

Ng, C. S. M., & Ng, S. S. L. (2022). Impact of the COVID-19 pandemic on children's mental health: A systematic review. *Frontiers in Psychiatry*, *13*, 975936.

Rones, M. & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical Child and Family Psychology Review, 3*(4), 223–241.

Twenge, J.M., Cooper, A.B., Joiner, T.E., Duffy, M.E., & Binau, S.G. (2019). Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005-2017. *Journal of Abnormal Psychology, 128*(3), 185-199.

U.S. Centers for Disease Control and Prevention (CDC) (2019). *Data and statistics on children's mental health. data and statistics*. Retrieved from https://www.cdc.gov/childrensmentalhealth/data.html

Zhou, X., Hetrick, S. E., Cuijpers, P., Qin, B., Barth, J., Whittington, C. J., Cohen, D., Del Giovane, C., Liu, Y., Michael, K. D., Zhang, Y., Weisz, J. R., & Xie, P. (2015). Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: A systematic review and network meta-analysis. *World Psychiatry*, *14*, 207-222.

Zhou, X., Zhang, Y., Furukawa, T. A., Cuijpers, P., Pu, J., Weisz, J. R., … & Xie, P. (2018). Different types and acceptability of psychotherapies for acute anxiety disorders in children and adolescents: A network meta-analysis. *JAMA Psychiatry*, *76*(1), 41–50.