ANALYSIS OF DOMAIN CHANGE IN THE INTEGRATED DEVELOPMENTAL EVALUATION AND ASSESSMENT (IDEA)

Washington State Department of CHILDREN, YOUTH & FAMILIES

Washington State Department of CHILDREN, YOUTH & FAMILIES

Date: January 26, 2026

OIAA | Prepared by Ian Edelstein | Approved for distribution by Catherine M Wilson, Evaluation & Research Administrator

ANALYSIS OF DOMAIN CHANGE IN THE INTEGRATED DEVELOPMENTAL EVALUATION AND ASSESSMENT (IDEA)

Introduction

Background

Overview of the IDEA Initiative

The Integrated Developmental Evaluation and Assessment (IDEA) was implemented by Juvenile Rehabilitation (JR) in 2022 and replaced the Integrated Treatment Assessment. The IDEA is based upon the Positive Achievement Change Tool[1] and is implemented by Washington state courts as the Juvenile Court Assessment Tool (JCAT). The initial IDEA contains 121 items, while IDEA reassessments consist of 82 of those questions.

A major update in the JCAT/IDEA was the introduction of 8 dynamic "need domains":

- School
- Associations
- Family
- Drugs/Alcohol
- Mental Health
- Attitudes/Behaviors
- Aggression
- Social Skills

These need domains were tested and refined with help from panels of experts. According to the authors, the domains allow for reassessment and evaluation of changes in need levels and treatment response. This aligns with Risk-Need-Responsivity principles.[2]

Prior OIAA reports on the IDEA have focused on factor analysis[3] of the eight need domains and analysis of association between all items on the IDEA and future recidivism. This study takes a different approach by looking at how JR clients' scores change over time and what factors might influence those changes.

[1] Hamilton et al. (2020). PACT Validation and Weighting Results Technical Report Deliverable 1: Updated PACT Risk and Needs Assessment. Washington State University. Department of Criminal Justice and Criminology. Institute for Criminal Justice.
[2] Bonta, J., & Andrews, D. A. (2007). Risk-need-responsivity model for offender assessment and rehabilitation. *Rehabilitation*, *6*(1), 1-22.
[3] https://dcyf.wa.gov/sites/default/files/pdf/reports/idea-factor-analysis-report.pdf

Goals of this report

This report answers four main questions:

1. How much do scores change in each of the 8 IDEA domains between assessments? (We look at changes from the first assessment to the second, and from the second to the third)?[4]
2. Do changes differ for different groups of youth (by gender, race, age, or facility type)?
3. Do different assessors/interviewers produce different results? (This could suggest bias or indicate a need for better training and quality control)
4. Which factors (like age, race, gender, facility, and time between assessments) are most strongly connected to score changes in each domain?

Key findings

Overall patterns of change

Most domains showed improvement (lower need scores) between the first and second assessment. In many cases, these improvements were significantly larger for females than males—even before accounting for other factors like age, race, or facility location.

However, between the second and third assessments, there was very little change on average across any domain or group. This could happen for several reasons:

- The biggest changes occur after intake/initial assessment when youth enter the JR system[5]. This could be due to structural changes with the move from the community to the institution.
- Some initial improvements may have been overestimated for certain groups (suggesting possible bias or inconsistent assessment quality).

What the detailed analysis revealed

We used statistical models to account for multiple factors at once: race, gender, age, days between assessments, facility location, and which staff member conducted the assessment. While many models struggled to explain changes in scores (possibly due to issues identified in previous research about how well these domains measure what they intend to), several important patterns emerged:

- **Assessor effects matter**: Even after accounting for everything else, which staff member conducted the assessment significantly affected scores in every model. A plausible explanation is that different assessors are rating youth differently in ways that reflect their own subjective judgment rather than actual changes in the youth. This needs more oversight and attention. These differences often align with extremely large (and possibly unrealistic) shifts in scores.

[4] Further change analysis, e.g. from second to third reassessment, or time 3 to 4, is not possible at this time due to low client numbers across groups.

[5] Ideally the IDEA is conducted within the first 30 days of intake, but the initial implementation required initial IDEAs to be developed for every JR client, some of whom were resident at JR for a year or more before the initial IDEA.

- **Facility differences**: Youth at Green Hill School consistently showed higher needs or worsening problems compared to other facilities.
- **Racial disparities**: For several domains (Attitudes/Behaviors, Aggression, and Mental Health), we found concerning patterns for Black/African American JR clients. Over time, they were assessed as having increasing behavioral and aggression needs but decreasing mental health needs. This may reflect bias in how assessors perceive and rate Black/African American clients and deserves further investigation.

Conclusions

For change scores to be meaningful and useful, assessments need to be:

- Conducted consistently by all assessors
- Done according to proper procedures
- Administered frequently enough to track real changes

As recommended in our previous report on factor analysis, JR needs to establish strong quality assurance systems, continuous improvement processes, and reliability checks (making sure different assessors rate the same youth similarly). These systems need to be in place long enough to build a new, improved dataset.

Based on this analysis, reassessments after the second time point (third assessment and beyond) appear to provide limited additional value under current practices. This doesn't mean all reassessment should stop, but there's no evidence here that continuing to measure all domains in later reassessments helps track meaningful change.

Technical Appendix

Methodology

Data and scoring

All initial and IDEA reassessments were pulled from the JR Automated Client Tracking (ACT) system for the period July 2022 (IDEA inception) through January 2025. A selection was then made of all first obligation initial IDEAs with at least two subsequent reassessments under the same obligation. After removing duplicate initial assessments, the final dataset contained 727 JR clients and a total of 2,183 IDEA assessments. There were 658 male clients in the dataset (90.5%) and 69 females (9.5%). The mean age at first assessment was 16.8 years with a range from 12 to 24 years.

Following the 2025 IDEA Factor Analysis report, the same continuous response coding was adopted for each assessment item with higher numbered responses indicating greater need, or risk. Domain scores were then constructed with only the items identified in the final factor models in the report. This approach was taken to attempt to reduce noise in the domain scores from items that did not contribute significantly to the same underlying latent construct (the need domain).

Static historical item scores (utilized in School, Family, Alcohol/Drug and Mental Health domains) were carried forward into the reassessments where those items are not reassessed. This ensures there is a score for each item in each assessment wave. Change scores were then built by subtracting the previous assessment's domain score from the new assessment's score. Following the recommendation of Labrecque et al.,[6] the raw difference or change score was then transformed into a percentage change score by dividing it by the previous score. For example, a one-point raw score change from 15 in time 1 to 16 in time 2 would result in a percentage change of 1/15= a 6.7% increase. The authors found that the "difference in effect size between the two change variables examined suggests that the use of the percentage change in scores affords a greater amount of incremental validity compared to that of the raw difference in scores." This means that small raw score changes in later reassessments can still be compared on the same scale with larger raw score changes common from the initial IDEA to the first reassessment.

Average change and distribution

Averaged, or mean, percentage change scores are calculated along with standard errors and 95% confidence intervals allowing for comparison across waves and between gender and racial groups. Essentially, when there is no overlap between the upper or lower end of two confidence intervals, those averages are statistically different. The spread of the confidence intervals is linked to the number of observations in the variable so smaller groups will have less precision and wider confidence intervals.

After exploring mean differences by wave, boxplots were generated to visualize the spread of the data and identify outlier scores that could influence the mean and affect further statistical analysis. In some statistical practices it is acceptable to remove outliers. In this study, where each change score is linked to two assessments, that approach would result in the removal of two data points and was not adopted.

[6] Labrecque, R. M., Smith, Paula, Lovins, B. K., & Latessa, E. J. (2014). The importance of reassessment: how changes in the LSI-R risk score can improve the prediction of recidivism. Journal of Offender Rehabilitation, 53(2), 116-128.

Mean scores were then compared across gender and racial groups. Multiracial groups were combined with their first identified category to reduce the presence of groups with numbers too small to report, e.g. Black/African American, Multiracial was coded as Black/African American. In some cases, small group numbers still limit interpretability.

Multivariate testing

In the last phase of analysis, linear regression models were built to include the available categorical and continuous independent variables. The domain change scores are assumed to be continuous in nature and suitable as dependent, or outcome, variables in regression analysis.

The categorical variables include gender, racial category, interaction effects between gender and race, institution at the time of the follow-up assessment, and assessment author at the time of follow-up[7].

Regression analysis requires that the intercept, or lowest level of each variable is excluded from the output reporting. As the largest group in number, White males from Echo Glen were the omitted, or reference category. The reference assessment author was one of the more productive, consistent scorers who had completed assessments throughout the duration of the data. Continuous variables include age at the time of the first assessment and days between the first and follow-up assessment. The reference value would be the lowest of each, e.g. 12 years at first assessment and the shortest duration between assessments.

Results

Attitude-Behavior Change

As shown in Table 1 below, average Attitude-Behavior change from the initial IDEA to the 1st reassessment (time 1 -time 2) is a reduction of 1.4% in need but is not statistically significant from 0 at a 95% confidence interval. Between 1st and 2nd reassessments (time 2-time 3), there was an average increase in Attitude-Behavior need of 3% and this change is significant from 0. It should be noted that only 339 cases included two reassessments (3 waves of data). The confidence intervals of the two means do not overlap, indicating that they are significantly different.

Table 1. Attitude/Behavior Domain change expressed as a percentage of previous score

Change Variable	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Att/Behav Time 1-2	-1.40	0.73	-2.83	0.04	727
Att/Behav Time 2-3	2.97	1.13	0.74	5.20	339

[7] Tests were conducted using more complex hierarchical and longitudinal models and including the location and author at the time of the first assessment in the change score. Due to client movements between institutions, author changes between waves, and the unbalanced nature of the panel data (i.e. small numbers of clients with as many as 9 assessment points), these more complex modeling techniques were not determined to be appropriate. Results did not differ significantly from the linear regression models presented and this approach allows for standardized coefficients and more interpretability.

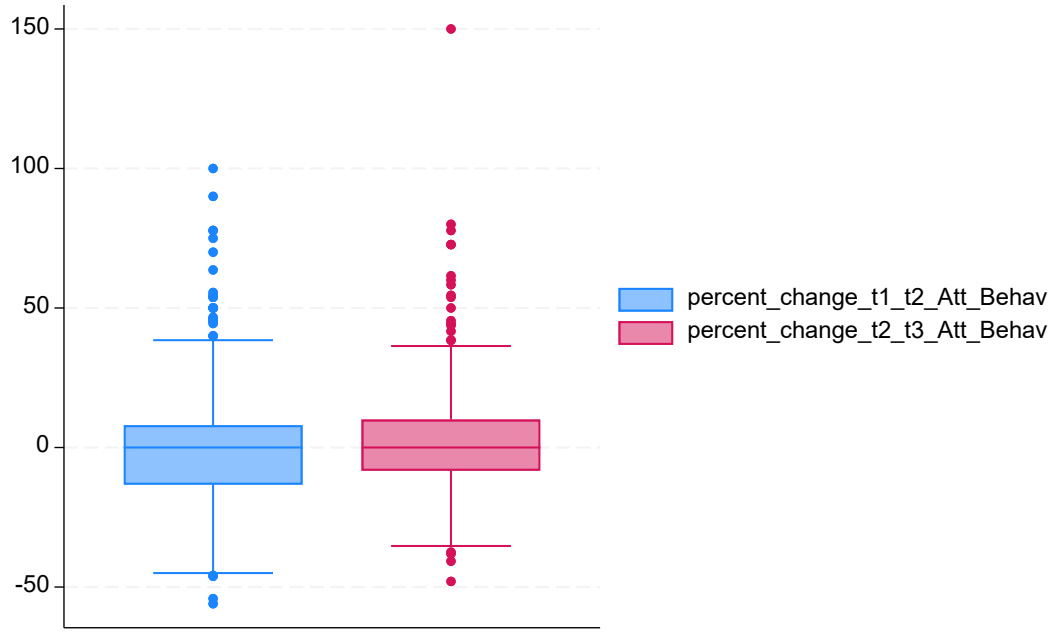
Distributions and outliers

Before exploring potential differences by gender or racial category, it is important to examine the overall spread of the data and the presence of potential outlier scores that could impact group differences. Boxplots are a straightforward way to visualize using five components:

- the median (or exact middle number in the variable scores)
- a box that contains the middle half of the scores, or the Interquartile Range (IQR)
- whiskers that show the spread of the upper and lower normal values (up to 1.5 times away from the median)
- outlier dots that show extreme scores that fall outside of the box and whiskers.

Figure 1 below shows the boxplots for Attitude-Behavior Domain change across the two time change periods. Overall, the narrow boxes with medians all near zero suggest that many of the scores represent minor changes in percentage terms. The whiskers are similar in nature, yet it is apparent that there are more negative scores (need reductions) from time 1-2 than in the subsequent wave. From Time 1 to Time 2, there were 27 outliers (3.7% of the total). From time 2 to time 3, there were 21 outliers (6.2% of the total).

Figure 1. Boxplots with outliers – Attitude/Behavior Domain change in percentage



Gender/Racial differences

Due to the small number of female clients in the dataset, it is only feasible to compare means across the first 3 waves of data, e.g. between initial IDEA and first reassessment and between first reassessment and second reassessment. It should be noted that there are only 20 female clients with data for these three waves compared to 319 male clients. As shown in Figure 2 below, there was a substantial drop in the average Attitude-Behavior score for females only from time 1-2 of 6.8%. This is significantly different than the changes seen across male groups as well as for females from time 2-3. The other gender and time change differences are not significantly different from one another, nor from a zero value.

Racial group differences

Within each wave (Tables 2 and 3 below), none of the racial group differences are statistically different from one another nor from a zero value. Graphs are not presented as there are no statistical differences to display.

Figure 2. Attitude/Behavior Domain change time 1-2 by gender with 95% confidence intervals

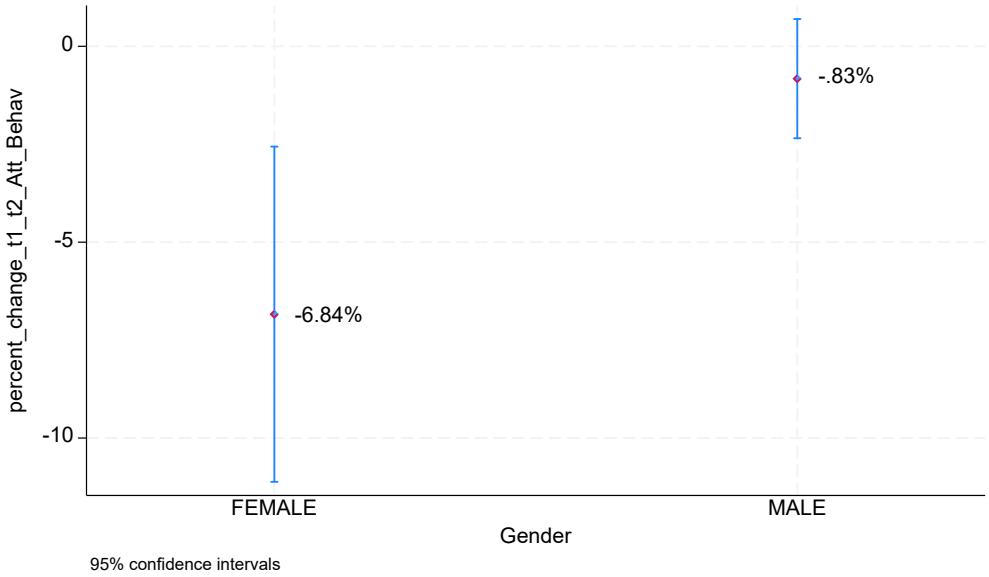


Figure 3. Attitude/Behavior Domain change time 2-3 by gender with 95% confidence intervals

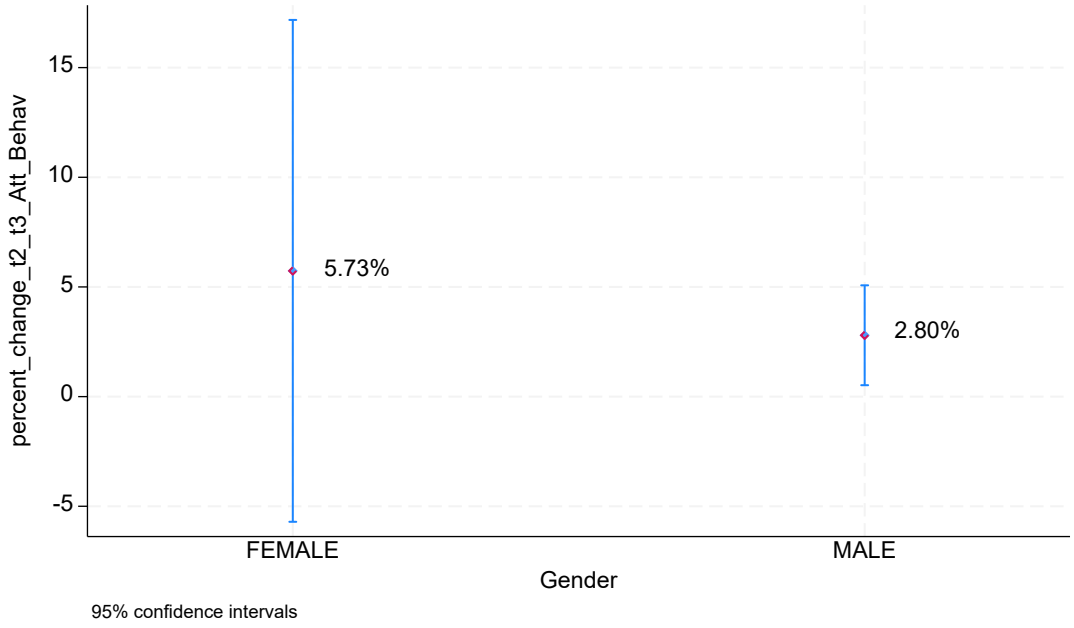


Table 2. Attitude/Behavior Domain change time 1-2 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	-6.84	2.15	-11.05	-2.63	69
Male	-0.83	0.77	-2.35	0.70	658
White	-3.30	1.20	-5.66	-0.95	275
Native	-5.53	2.56	-10.55	-0.51	40
Asian	-1.53	3.60	-8.61	5.55	32
Black	1.64	1.49	-1.28	4.57	190
Latino	-0.94	1.47	-3.82	1.93	175

Table 3. Attitude/Behavior Domain change time 2-3 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	5.73	5.47	-5.02	16.48	20
Male	2.75	1.16	0.52	5.07	319
White	1.87	1.66	-1.40	5.13	121
Native	-4.07	2.28	-8.56	0.43	15
Asian	4.93	6.90	-8.64	18.51	15
Black	3.58	2.32	-0.99	8.15	93
Latino	4.53	2.49	-0.38	9.43	89

Aggression Domain Change

Average change

As shown in Table 4 below, average Aggression change from time 1 -time 2 is a reduction of 4.9% in need and is statistically from 0 at a 95% confidence interval. Between time 2-time 3, there was an average increase in Attitude-Behavior need of 1.2% but this change is not significant from 0. It should be noted that only 339 cases included two reassessments.

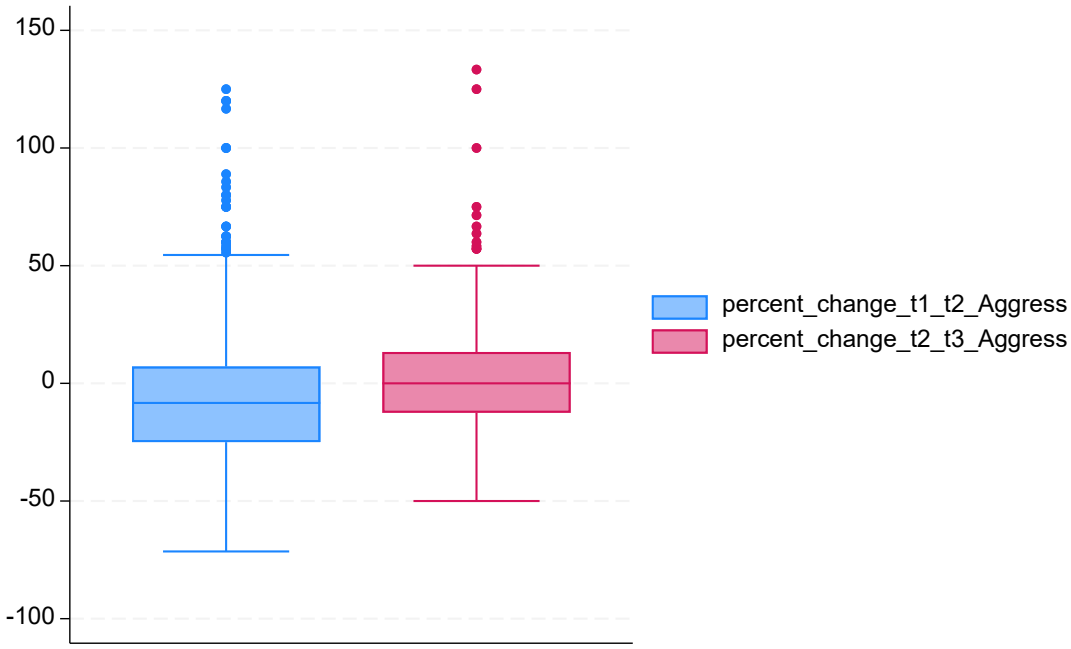
Table 4. Aggression Domain change expressed as a percentage of previous score

Change Variable	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Aggress. Time 1-2	-4.92	1.03	-6.93	-2.90	727
Aggress. Time 2-3	1.22	1.29	-1.32	3.76	339

Distributions and outliers

Figure 4 below shows the boxplots for Aggression Domain change across the two time change periods. The narrow boxes with medians near zero indicate that many of the scores represent minor changes in percentage terms. There are more negative scores (need reductions) from time 1-2 than in the subsequent wave. From Time 1 to Time 2, there were 26 outliers (3.6% of the total). From time 2 to time 3, there were 11 outliers (3.2% of the total). All of the outliers are in the direction of increased Aggression need.

Figure 4 Boxplots with outliers – Aggression Domain change in percentage



Gender and Racial differences

As shown in Figure 5 below there was a substantial drop in the average Aggression score for females from time 1-2, of 18.5%, significantly greater that the reduction for males of 3.5%. In time 2-3, both females and males saw a slight Aggression need increase (of 0.1% and 1.3%, respectively), though not statistically different from zero.

In Time 1-2 (Table 5 below), none of the Aggression difference by racial category are significantly different from one another (i.e. they have overlapping confidence intervals). However, White, Native American and Asian groups had means significantly lower than zero (Aggression reductions, on average) while Black and Latino groups had values that were not significantly different from zero. In Time 2-3 (Table 6 below), there are no nonzero values for Aggression change across racial groups. Some groups have low counts, and this coincides with larger standard errors.

Figure 5. Aggression Domain change time 1-2 by gender with 95% confidence intervals

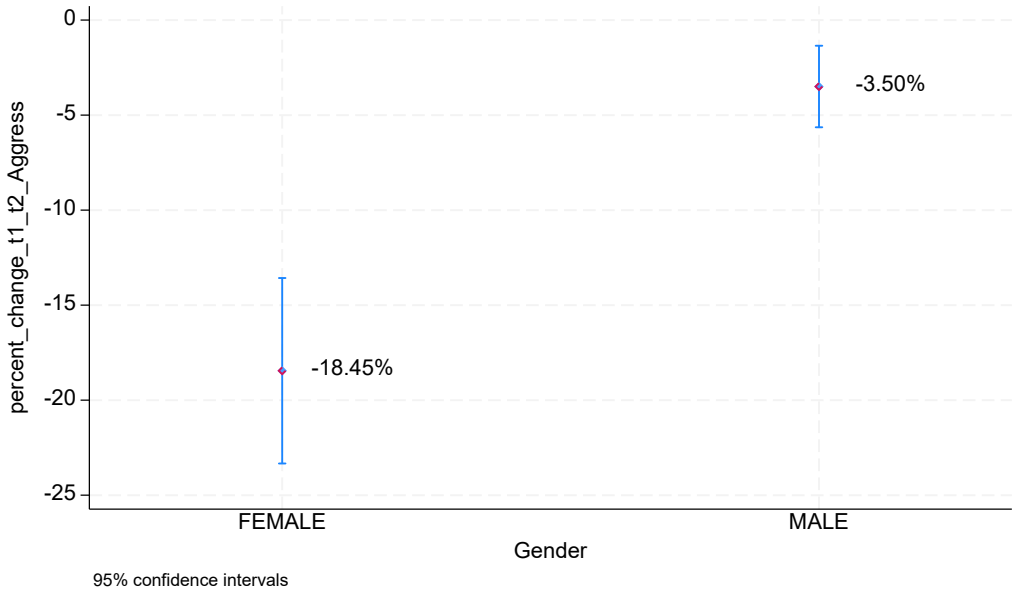


Figure 6. Aggression Domain change time 2-3 by gender with 95% confidence intervals

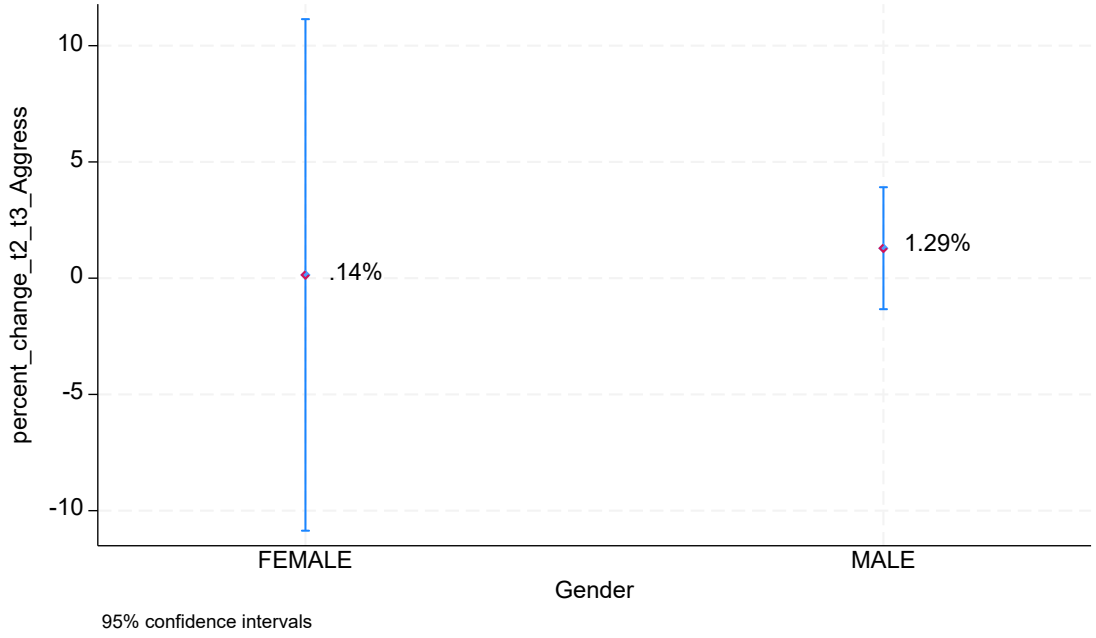


Table 5. Aggression Domain change time 1-2 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	-18.45	2.45	-23.25	-13.65	69
Male	-3.50	1.09	-5.64	-1.35	658
White	-7.08	1.61	-10.24	-3.91	275
Native	-10.81	2.79	-16.27	-5.34	40
Asian	-8.33	4.14	-16.47	-0.20	32
Black	-1.40	2.27	-5.85	3.06	190
Latino	-3.32	2.10	-7.44	0.79	175

Table 6. Aggression Domain change time 2-3 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	.14	5.26	-10.20	10.48	20
Male	1.29	1.33	-1.33	3.91	319
White	-0.34	1.82	-3.92	3.24	121
Native	2.81	4.84	-6.71	12.33	15
Asian	7.92	7.19	-6.22	22.07	15
Black	2.08	2.42	-2.68	6.84	93
Latino	-0.05	3.05	-6.04	5.94	89

Skills Domain Change

Average change

As shown in Table 7 below, average Social Skills change from the initial IDEA to the 1st reassessment is a reduction of 9.1% in need and is statistically significant from 0 at a 95% confidence interval. Between 1st and 2nd reassessments (time 2-time 3), there was an average decrease in Skills domain need of 1.2%, not significantly different from 0. The lack of overlap in confidence intervals indicates that the Skills change in time 1-2 is significantly greater than time 2-3.

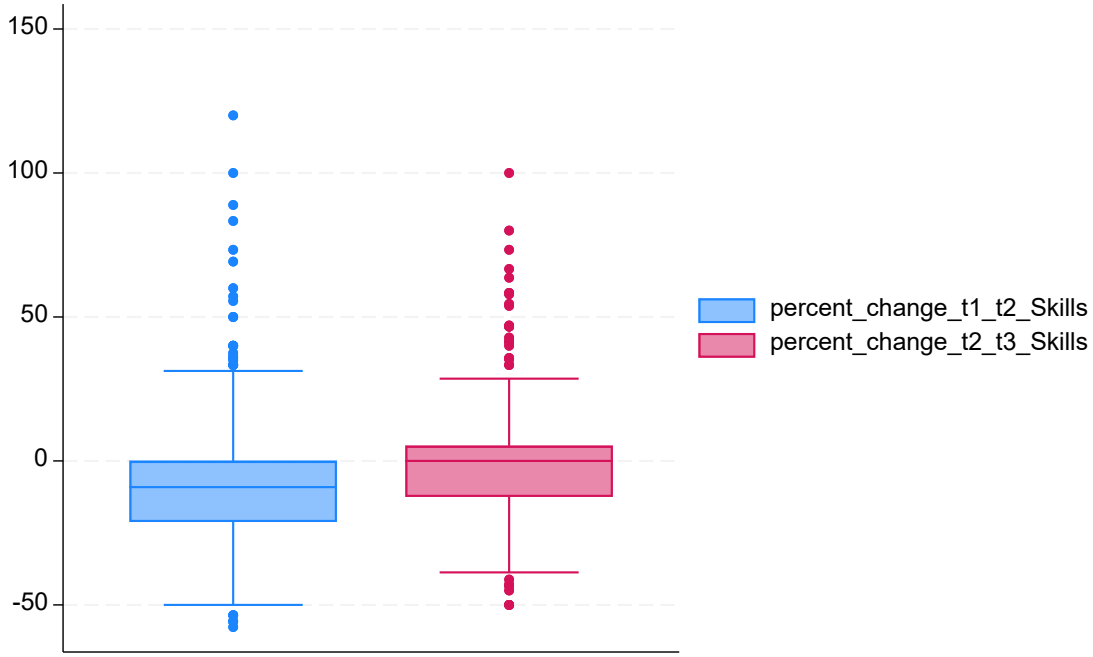
Table 7. Skills Domain change expressed as a percentage of previous score

Change Variable	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Skills Time 1-2	-9.10	0.72	-10.51	-7.68	727
Skills Time 2-3	-1.16	1.13	-3.39	1.06	339

Distributions and outliers

Figure 7 below shows the boxplots for Skills Domain change across the two time change periods. There are more negative scores (need reductions) from time 1-2 than time 2-3. From Time 1 to Time 2, there were 26 outliers (3.6% of the total). From time 2 to time 3, there were 30 outliers (8.8% of the total), primarily in the Skill Need increase direction. This could be an area where stronger quality assurance practices are warranted.

Figure 7. Boxplots with outliers – Skills Domain change in percentage



Gender and racial group differences

As shown in Figure 8 and Table 8 below, there was a substantial drop in the average Skills Domain score for females from time 1-2, of 16.5%, significantly greater that the reduction for males of 8.3%. In time 2-3, females saw an increase of 7.5%, while males saw a slight Skills Domain decrease of 1.7%, though neither are statistically different from zero.

Figure 8. Skills Domain change time 1-2 by gender with 95% confidence intervals

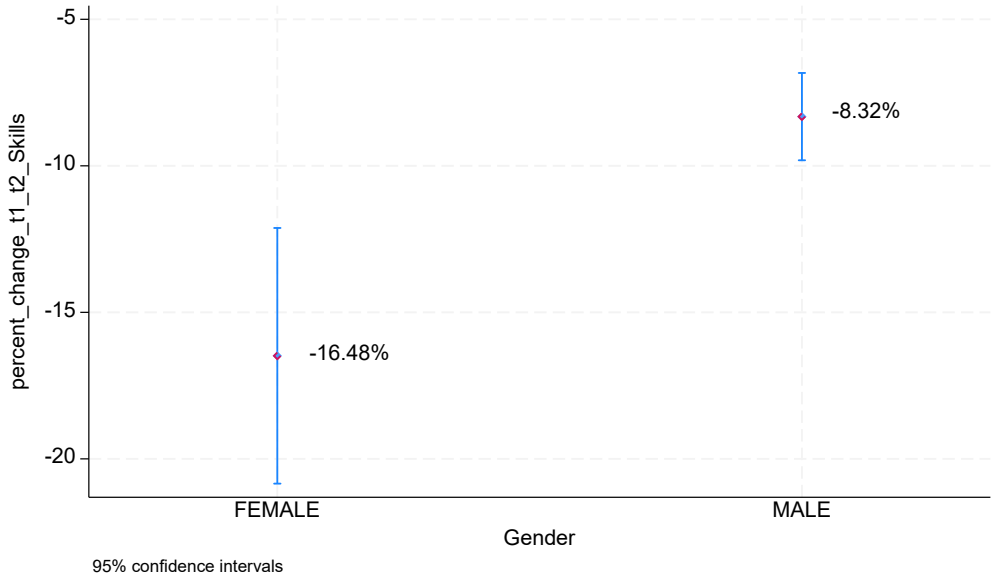
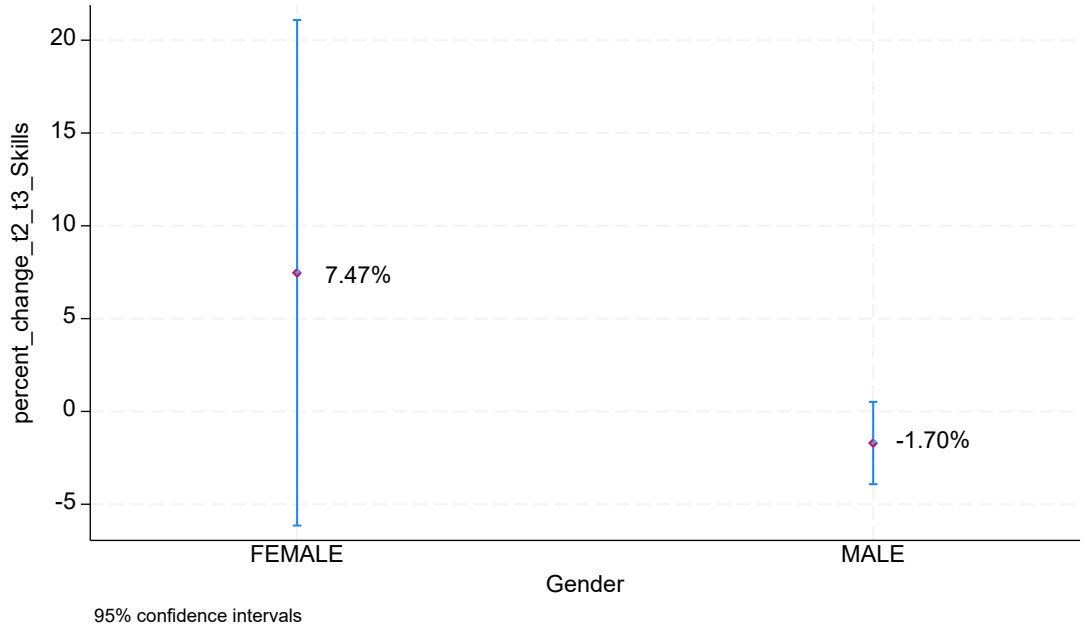


Figure 9. Skills Domain change time 2-3 by gender with 95% confidence intervals



Across racial groups, there were no significant Skills domain change differences in time 1-2 (Table 8 below) and all are negative and nonzero. In time 2-3 (Table 9 below), only the Native American group saw a nonzero, negative Skills change score, but this should be interpreted with caution due to the small group size (n=15) and the presence of outliers.

Table 8. Skills Domain change time 1-2 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	-16.48	2.19	-20.78	-12.19	69
Male	-8.32	0.76	-9.81	-6.83	658
White	-9.66	1.31	-12.23	-7.08	275
Native	-10.84	2.96	-16.65	-5.03	40
Asian	-10.29	2.61	-15.42	-5.17	32
Black	-7.24	1.44	-10.07	-4.42	190
Latino	-9.66	1.25	-12.12	-7.20	175

Table 9. Skills Domain change time 2-3 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	7.47	6.51	-5.33	20.27	20
Male	-1.70	1.13	-3.92	0.51	319
White	-3.27	1.86	-6.93	0.40	121
Native	-9.71	3.98	-17.54	-1.89	15
Asian	0.61	6.57	-12.31	13.53	15
Black	1.82	2.23	-2.56	6.20	93
Latino	0.08	2.19	-4.23	4.39	89

School Domain Change

Average change

As shown in Table 10 below, average School Domain change from the initial IDEA to the 1st reassessment is a reduction of 9.8% in need and is statistically significant from 0. However, there are only 249 nonmissing values in this variable due to many not applicable responses that could not be used to calculate a domain change score. Between 1st and 2nd reassessments, there was an average increase in School need of 0.2%, though not significant from 0. It should be noted that only 105 cases included two reassessments, again due to the frequency of not applicable School domain responses. The difference between the two mean scores is significant.

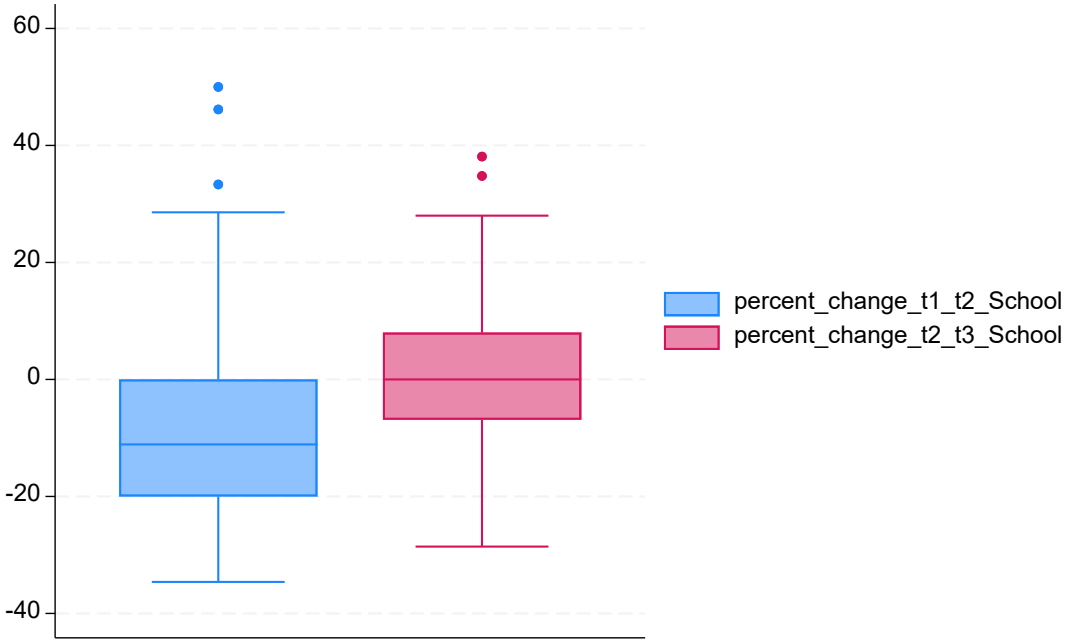
Table 10. School Domain change expressed as a percentage of previous score

Change Variable	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
School Time 1-2	-9.80	0.90	-11.58	-8.02	249
School Time 2-3	0.24	1.21	-2.16	2.65	105

Distributions and outliers

Figure 10 below shows the boxplots for School Domain change across the two time change periods. Compared to other boxplots, there appears to be a greater spread of the boxes containing the middle 50% of scores. There are also fewer outliers than in other domain change variables. From Time 1 to Time 2, there were 3 outliers (1.2% of the total). From time 2 to time 3, there were also 2 outliers (1.9% of the total).

Figure 10. Boxplots with outliers – School Domain change in percentage



Gender and Racial group differences

As shown in Figure 11 and Table 11 below, there was a substantial drop in the average School domain score for females from time 1-2, of 14.3%, though not significantly greater that the reduction for males of 9.2%. In time 2-3 (Figure 12 and Table 12 below), females saw a decrease of 5%, while males saw a slight School domain increase of 0.6%, though neither are statistically different from zero.

Figure 11. School Domain change time 1-2 by gender with 95% confidence intervals

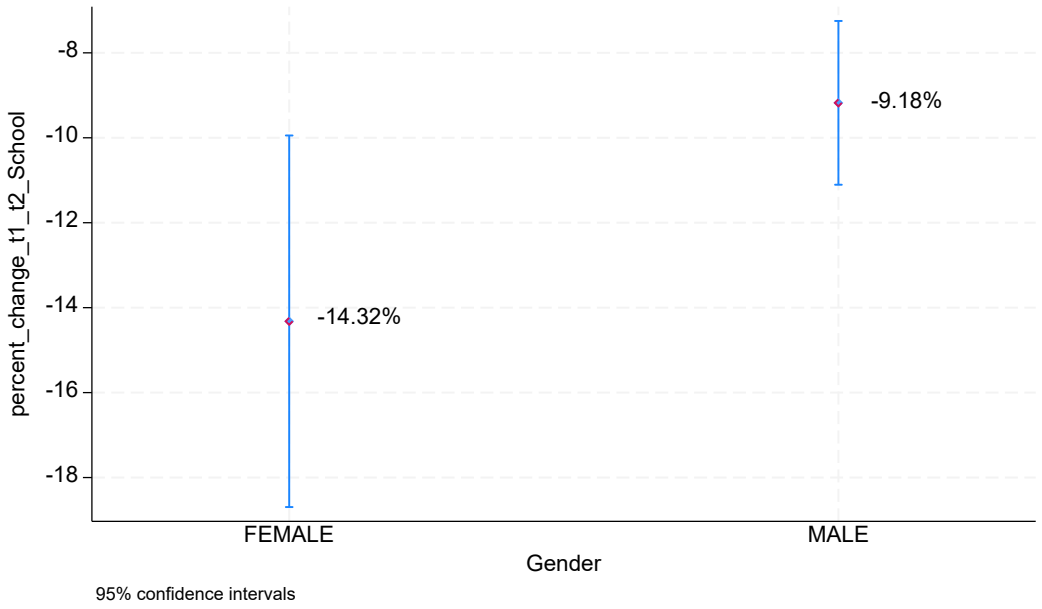
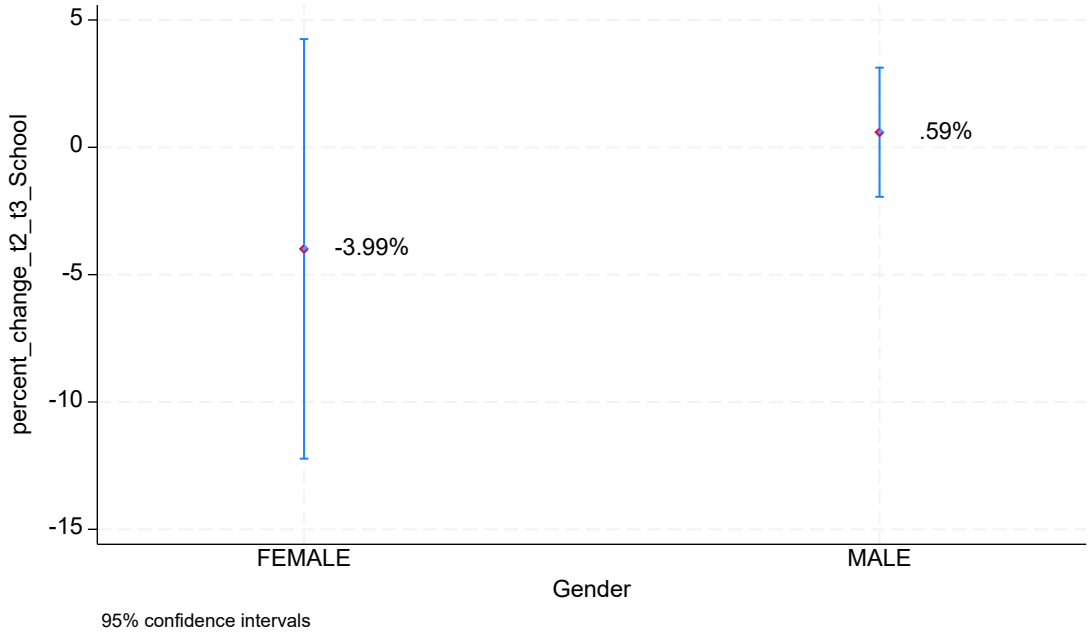


Figure 12. School Domain change time 2-3 by gender with 95% confidence intervals



School domain racial group differences were not significantly different from one another and were all negative in time 1-2 (Table 11 below). Native American was the only group with a mean School change from time 1-2 that was not significantly different from zero. In time 2-3 (Table 12 below), racial group differences are again not significant and nonmissing values are too few for Native American and Asian groups to allow for meaningful comparison.

Table 11. School Domain change time 1-2 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	-14.32	2.13	-18.53	-10.11	30
Male	-9.18	0.98	-11.10	-7.25	219
White	-12.25	1.29	-14.78	-9.72	92
Native	-6.20	6.19	-18.40	6.00	15
Asian	-14.21	3.19	-20.50	-7.92	13
Black	-6.34	1.76	-9.80	-2.88	66
Latino	-10.52	1.75	-13.97	-7.08	59

Table 12. School Domain change time 2-3 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	7.47	6.51	-5.33	20.27	8
Male	-1.70	1.13	-3.92	0.51	97
White	-3.27	1.86	-6.93	0.40	34
Native	-9.71	3.98	-17.54	-1.89	8
Asian	0.61	6.57	-12.31	13.53	5
Black	1.82	2.23	-2.56	6.20	31
Latino	0.08	2.19	-4.23	4.39	27

Family Domain Change

Average change

As shown in Table 13 below, average Family domain change from time 1 -time 2 is a reduction of 7.3% in need, statistically different from 0. However, there were only 319 nonmissing values due to not applicable responses in the Family domain. Between time 2-time 3, there was an average increase in Family need of 1.1%, though not significant from 0 (and included only 91 complete responses). The confidence intervals of the two means do not overlap and are significantly different.

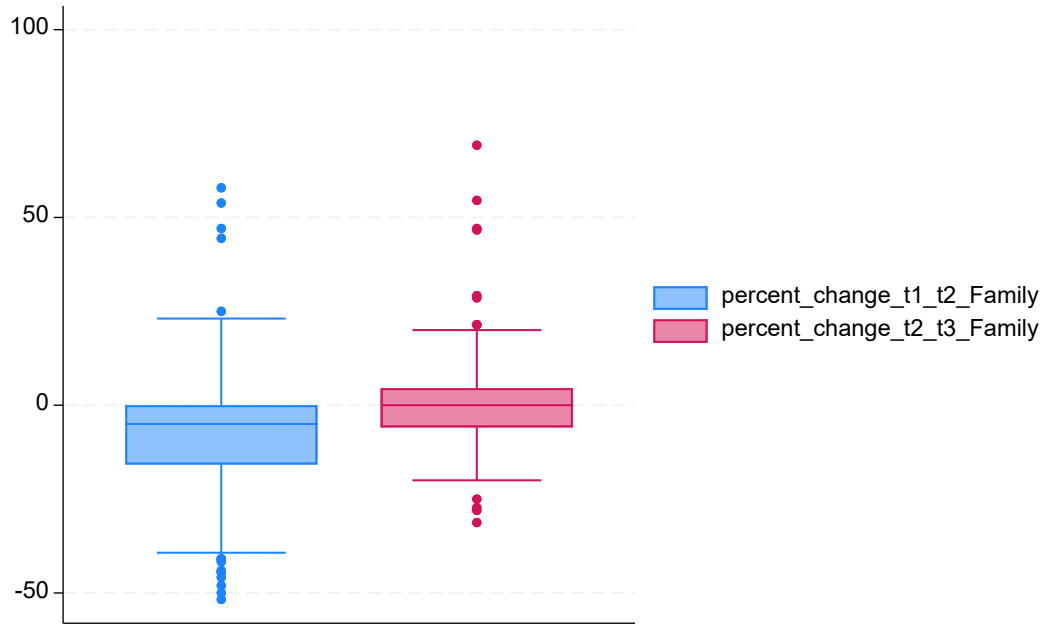
Table 13. Family Domain change expressed as a percentage of previous score

Change Variable	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Family Time 1-2	-7.28	0.87	-9.00	-5.57	319
Family Time 2-3	1.05	1.67	-2.26	4.37	91

Distributions and outliers

Figure 13 below shows the boxplots for Family domain change across the two time change periods. From Time 1 to Time 2, there were 16 outliers (5.02% of the total). From time 2 to time 3, there were 12 outliers (13.19% of the total).

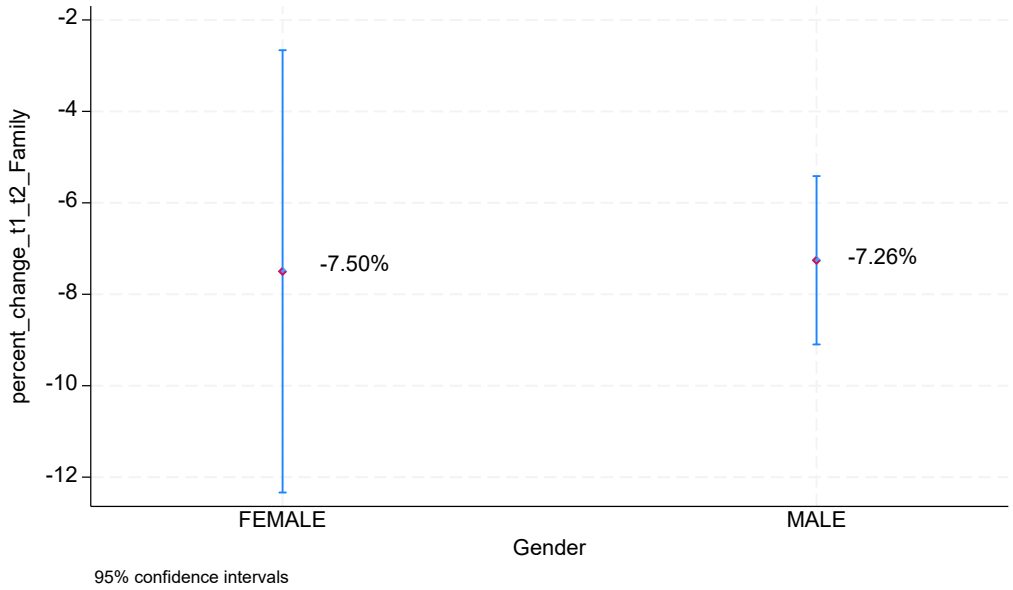
Figure 13. Boxplots with outliers – Family Domain change in percentage



Gender and racial group differences

As shown in Figure 14 and Tables 14 and 15 below, there were substantial drops in the average Family domain score for both females and males from time 1-2, of 7.5% and 7.3%, respectively. In time 2-3 (Table 15 below), female nonmissing values are too few for comparison. Thus, a figure for Family time 2-3 is not presented.

Figure 14. Family Domain change time 1-2 by gender with 95% confidence intervals



In the Family Domain, none of the changes are significantly different across racial groups in either wave (Tables 14 and 15 below). However, some subgroup values are significantly less than zero in time 1-2 (all groups except Asian, though nonmissing values are only 11 cases). In time 2-3, none of the change scores are significantly different from zero and female, Native American and Asian subgroups all have inadequate numbers for comparison.

Table 14. Family Domain change time 1-2 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	-7.50	2.38	-12.19	-2.81	36
Male	-7.26	0.94	-9.10	-5.42	283
White	-9.57	1.35	-12.22	-6.92	108
Native	-7.80	3.55	-14.79	-0.81	15
Asian	-5.36	2.89	-11.05	0.33	11
Black	-6.36	1.72	-9.75	-2.98	70
Latino	-5.25	1.94	-9.07	-1.43	74

Table 15. Family Domain change time 2-3 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	8.82	6.45	-3.99	21.64	7
Male	0.41	1.72	-3.01	3.82	84
White	-1.13	2.99	-7.07	4.81	39
Native	5.93	6.53	-7.05	18.90	8
Asian	-4.44	4.44	-13.28	4.39	3
Black	2.08	2.36	-2.61	6.77	19
Latino	3.15	3.29	-3.38	9.69	21

Associations Domain Change

Average change

As shown in Table 16 below, average Associations Domain change from time 1 -time 2 is a reduction of 9.6% in need and is statistically significant from 0. Between time 2-time 3, there was an average increase in Associations need of 0.7% but not significant from 0. The means are not statistically equal.

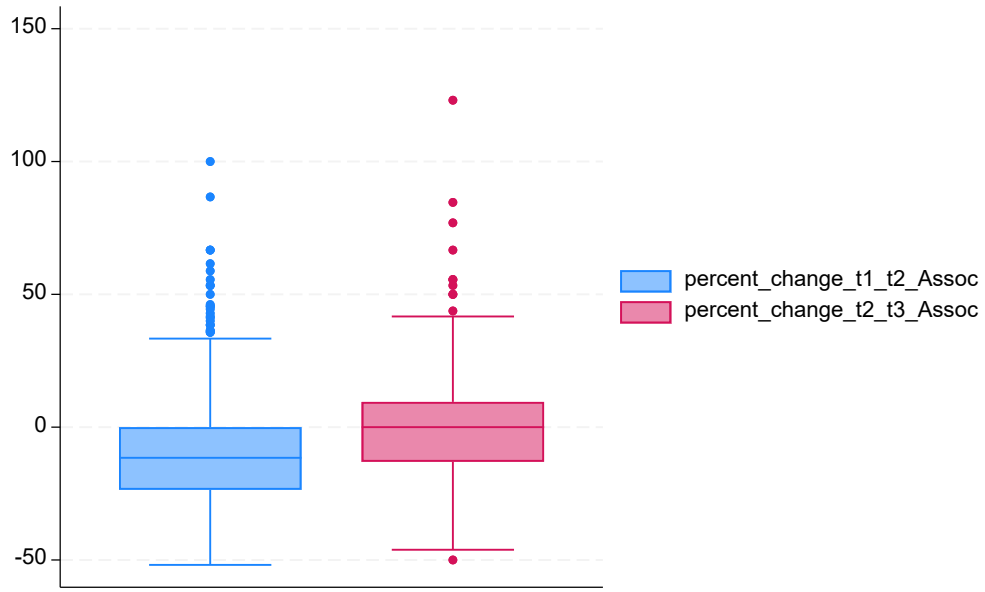
Table 16. Associations Domain change expressed as a percentage of previous score

Change Variable	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Assoc. Time 1-2	-9.64	0.75	-11.11	-8.17	720
Assoc. Time 2-3	0.07	1.17	-2.24	2.37	338

Distributions and outliers

Figure 15 below shows the boxplots for Associations domain change across the two time change periods. Overall, the narrow boxes with medians all near zero suggest that many of the scores represent minor changes in percentage terms. The boxes and whiskers are similar in size, yet it is apparent that there are more negative scores (need reductions) from time 1-2 than time 2-3. From Time 1 to Time 2, there were 24 outliers (3.33% of the total). From time 2 to time 3, there were 11 outliers (3.25% of the total).

Figure 15. Boxplots with outliers – Associations Domain change in percentage



Gender and racial group differences

As shown in Figure 16 and Table 17 below, there were substantial drops in the average Associations need domain score for both females and males from time 1-2, of 17.5% and 8.8%, respectively, significantly more change for females than males. In time 2-3 (Figure 17 and Table 18 below), values are both not significantly different from zero.

Figure 16. Associations Domain change time 1-2 by gender with 95% confidence intervals

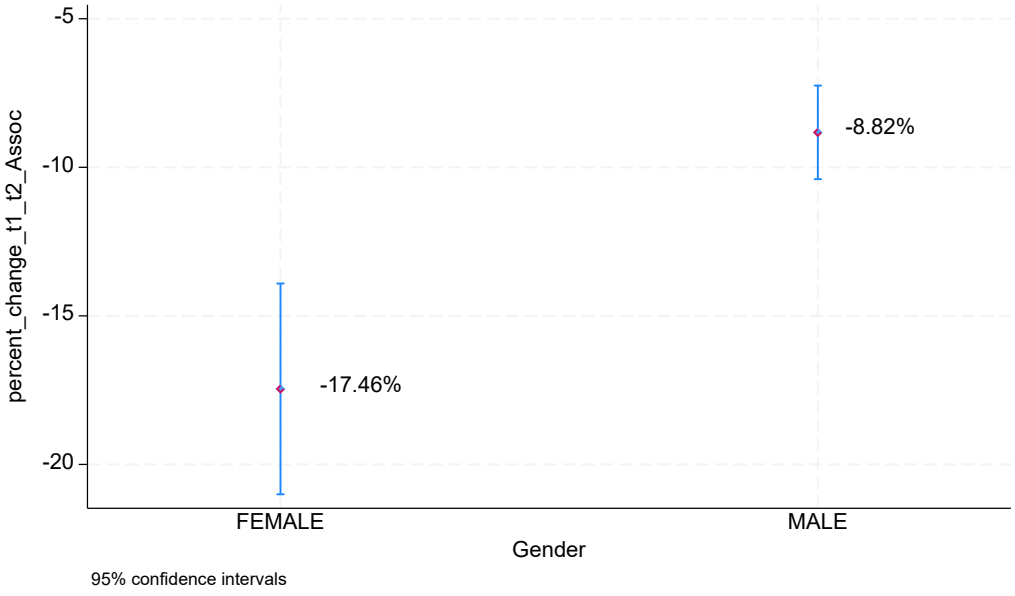
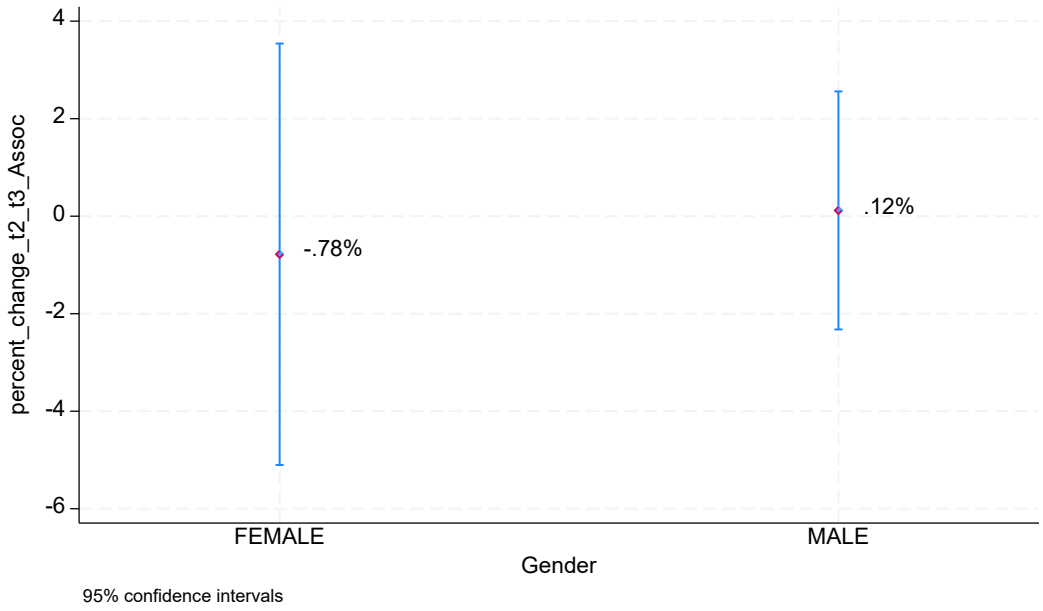


Figure 17. Associations Domain change time 2-3 by gender with 95% confidence intervals



In time 1-2 (Table 17 below), racial subgroup differences are not significant, yet all means are nonzero and in the negative direction (need reduction). In time 2-3 (Table 18 below), none of the Association domain changes by racial subgroup are significantly different from zero.

Table 17. Associations Domain change time 1-2 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	-17.46	1.78	-20.95	-13.97	68
Male	-8.82	0.80	-10.40	-7.25	652
White	-10.11	1.26	-12.58	-7.63	273
Native	-9.27	3.28	-15.71	-2.83	37
Asian	-12.20	3.04	-18.17	-6.22	32
Black	-9.61	1.51	-12.57	-6.65	188
Latino	-9.85	1.28	-12.36	-7.35	175

Table 18. Associations Domain change time 2-3 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	-0.78	2.07	-4.84	3.28	20
Male	0.11	1.24	-2.32	2.56	318
White	2.36	2.29	-2.14	6.87	121
Native	-3.73	4.05	-11.70	4.24	15
Asian	2.65	5.09	-7.37	12.67	15
Black	-2.78	2.00	-6.70	1.15	93
Latino	1.46	2.00	-2.48	5.40	88

Alcohol/Drugs Domain Change

Average change

As shown in Table 19 below, average Alcohol/Drugs change from the initial IDEA to the 1st reassessment is a reduction of 13.8% in need, statistically greater than 0. Between 1st and 2nd reassessments, there was an average increase in Alcohol/Drugs need of 3.7% and this change is also significant from 0. This is the first domain where overall change scores are significant and run in opposite directions. It should be noted that the Alcohol/Drugs domain did not exhibit adequate performance in factor analysis and thus may suffer from measurement error, or noise. It is also possible that increases in alcohol/drug need are aligned with positive drug screenings as youth gain access to such substances while in JR residency.

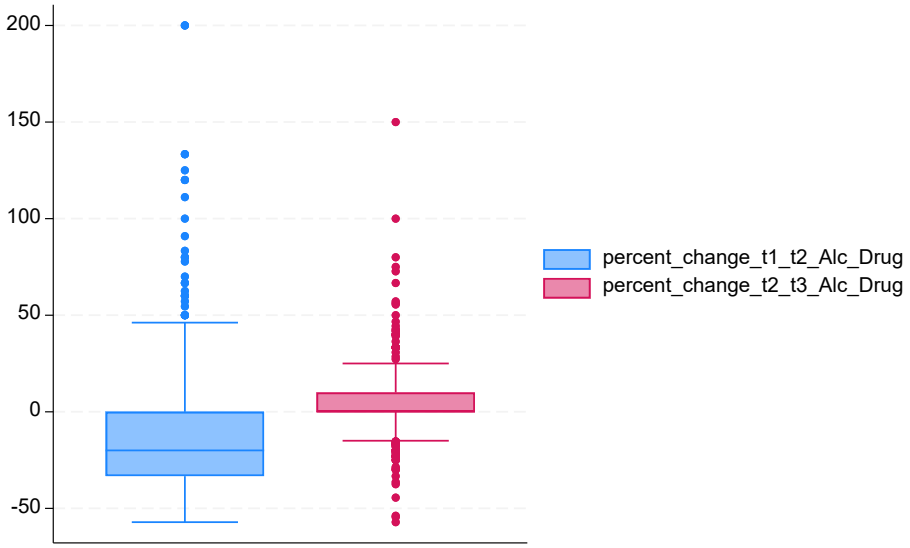
Table 19. Alcohol/Drugs Domain change expressed as a percentage of previous score

Change Variable	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Alc/Drugs Time 1-2	-13.84	1.08	-15.95	-11.72	727
Alc/Drugs Time 2-3	3.67	1.12	1.48	5.87	339

Distributions and outliers

Figure 18 below shows the boxplots for Alcohol/Drugs domain change across the two time change periods. The difference in box and whisker sizes demonstrate there were more negative, nonzero scores in time 1-2. In time 2-3, there were substantial numbers with no change, but also a large quantity of outliers. From Time 1 to Time 2, there were 27 outliers (3.71% of the total). From time 2 to time 3, there were 75 outliers (22.12% of the total).

Figure 18. Boxplots with outliers – Alcohol/Drugs Domain change in percentage



Gender and racial differences

As shown in Figure 19 below there were substantial - and statistically different - drops in the average Alcohol-Drug domain score from time 1-2 for females (-26.7%) and males (-12.5%). As seen in the overall mean change above, from time 2-3 (Figure 20 below), male scores increased in need by 3.8% (significantly greater than zero), while female scores increased by 1% (not significant from zero).

Figure 19. Alcohol-Drug Domain change time 1-2 by gender with 95% confidence intervals

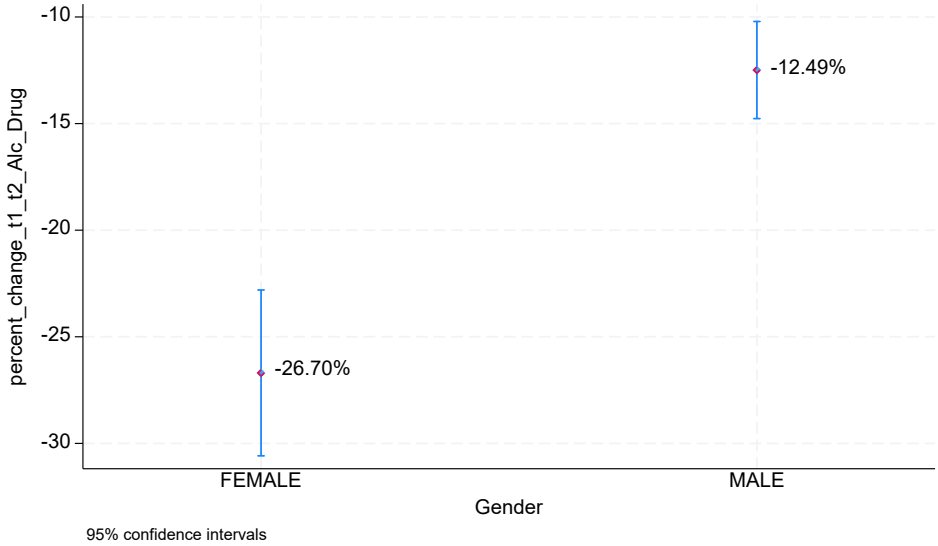
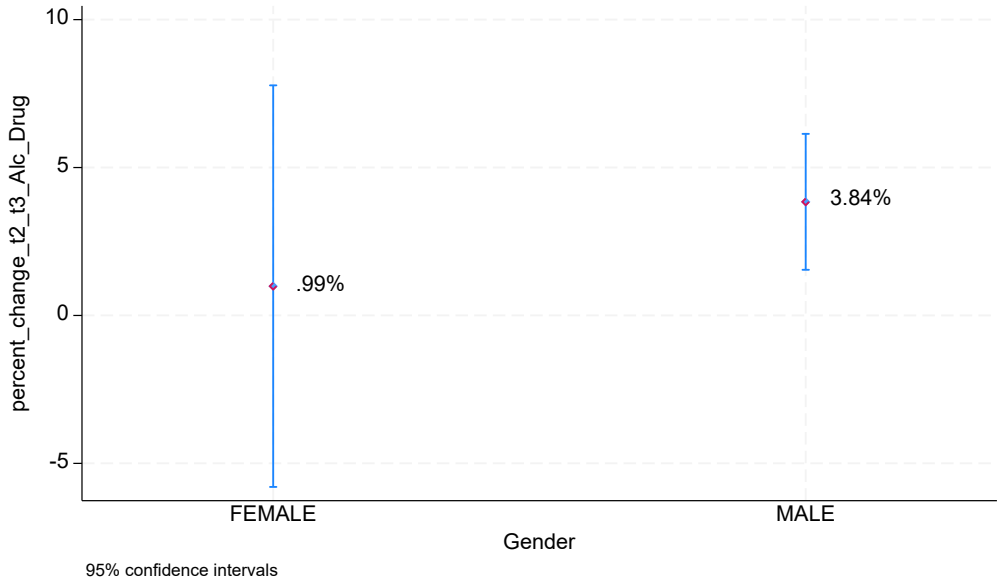


Figure 20. Alcohol-Drug Domain change time 2-3 by gender with 95% confidence intervals



There are no significant differences across racial groups in either wave (Tables 20 and 21, below). All groups have negative, nonzero means for time 1-2. Only Blacks had a nonzero increase in Alcohol/Drug needs in time 2-3.

Table 20. Alcohol/Drug Domain change time 1-2 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	-26.70	1.95	-30.52	-22.87	68
Male	-12.49	1.16	-14.77	-10.21	658
White	-16.25	1.46	-19.11	-13.39	275
Native	-18.22	4.13	-26.35	-10.10	40
Asian	-18.22	2.99	-24.08	-12.35	32
Black	-8.73	2.64	-13.90	-2.55	190
Latino	-15.43	1.85	-19.07	-11.80	175

Table 21. Alcohol/Drug Domain change time 2-3 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	0.99	3.24	-5.39	7.37	20
Male	3.84	1.17	1.54	6.14	319
White	2.15	1.46	-0.73	5.02	121
Native	-1.87	4.15	-10.04	6.30	15
Asian	12.19	10.25	-7.97	32.36	15
Black	6.01	2.52	1.06	10.97	93
Latino	3.28	1.88	-0.43	6.98	89

Mental Health Domain Change

Average change

As shown in Table 22 below, average Mental Health change from time 1-time 2 is a reduction of 0.3% in need but is not statistically significant from 0. Between time 2-time 3, there was an average increase in Mental Health need of 1% but is again not statistically significant from 0. The poor factor analysis results for this domain could influence the change scoring.

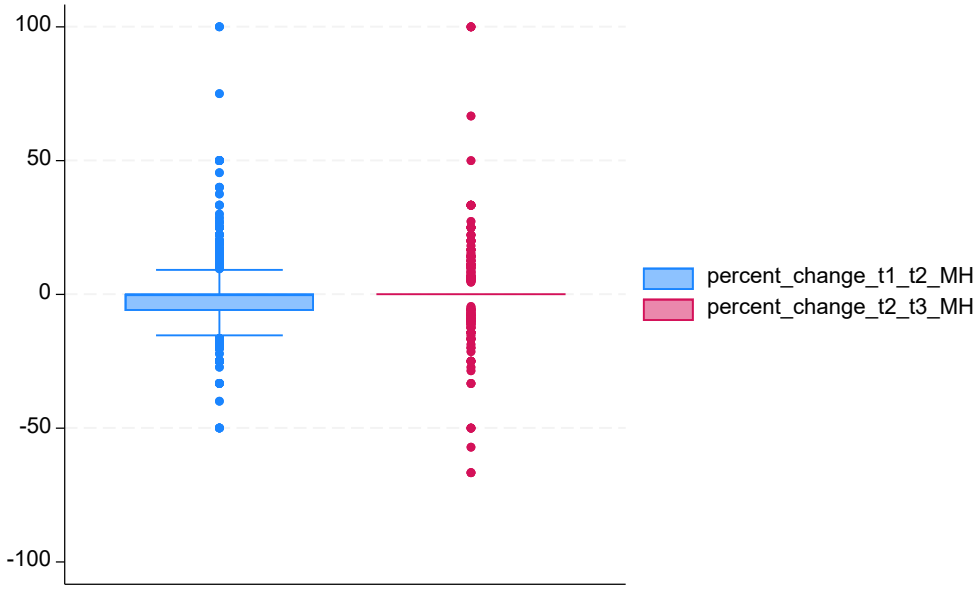
Table 22. Mental Health Domain change expressed as a percentage of previous score

Change Variable	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Mental Health Time 1-2	-0.34	0.50	-1.33	0.65	727
Mental Health Time 2-3	1.01	0.89	-0.74	2.76	339

Distributions and outliers

The boxplots for Mental Health Domain change across the two time change periods demonstrate rather problematic distributions with little variation for the interquartile and many outliers (Figure 21 below). From Time 1 to Time 2, there were 161 outliers (22.15% of the total). From time 2 to time 3, the interquartile range is 0, making it impossible to accurately identify outliers (effectively any non-zero value becomes an outlier).

Figure 21. Boxplots with outliers – Mental Health Domain change in percentage



Gender and racial group differences

As shown in Figures 22 and 23 and Tables 23 and 24 below, gender differences are not significant nor are mean Mental Health domain change scores different from zero.

Figure 22. Mental Health Domain change time 1-2 by gender with 95% confidence intervals

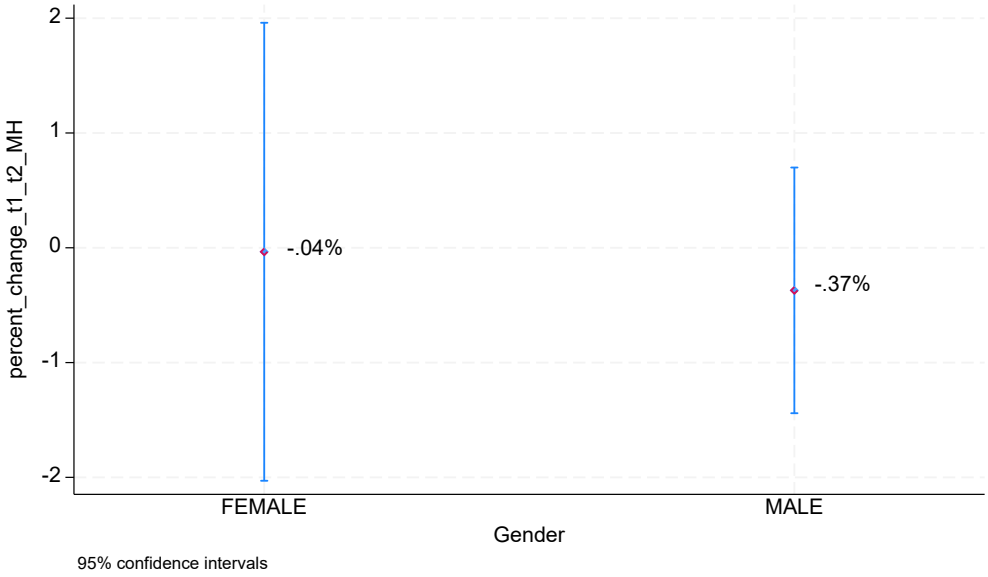
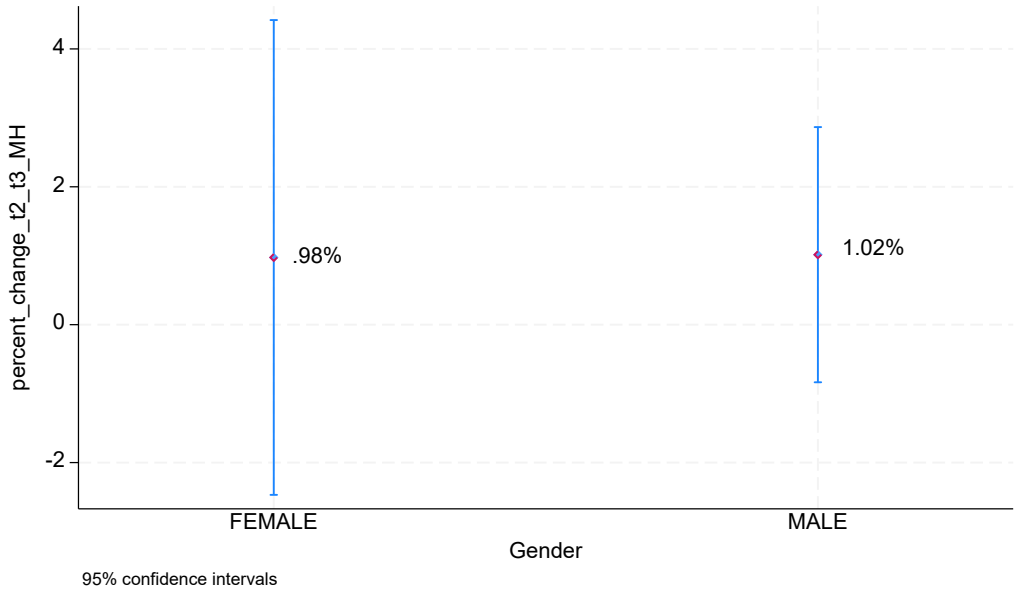


Figure 23. Mental Health Domain change time 2-3 by gender with 95% confidence intervals



As with the lack of gender differences, racial group differences in the Mental Health change scores are not statistically different from one another in either time 1-2 or time 2-3. However, the mean reduction for Black persons from time 1-2 of 2.5% is statistically different from zero.

Table 23. Mental Health Domain change time 1-2 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	-0.04	1.00	-2.00	1.93	68
Male	-0.37	0.55	-1.44	0.70	658
White	0.97	0.87	-0.74	1.69	275
Native	-3.33	2.02	-7.30	0.64	40
Asian	-0.06	1.66	-3.32	3.21	32
Black	-2.48	0.88	-4.21	-0.75	190
Latino	0.46	0.99	-1.48	2.39	175

Table 24. Mental Health Domain change time 2-3 by subgroup

Gender/racial group	Mean	Standard error	95% Confidence interval (lower end)	95% Confidence interval (upper end)	Number of cases with nonmissing values
Female	0.98	1.65	-2.26	4.21	20
Male	1.02	0.94	-0.84	2.87	319
White	0.23	1.29	-2.31	2.77	121
Native	2.11	2.51	-2.83	7.05	15
Asian	0.51	4.91	-9.14	10.16	15
Black	3.25	1.77	-0.23	6.73	93
Latino	0.06	2.00	-3.88	3.99	89

Multivariate analysis

Separate multivariate linear regression models were run for each of the need domain change scores from time 1-2 and time 2-3 (these are known as the outcome-or dependent-variables). This method yields an overall percentage of variance explained by each model[8] along with standardized coefficients for each independent variable and a p-value or significance level. Coefficients, in these models, represent the difference in percentage point change from the reference category, all else being equal.

The Table 25 below summarizes those independent variables that explain significant variance in the outcome or dependent variable (each need domain and change score period).[9] The independent variables include gender and racial group effects, interaction between gender and race, age at the earlier assessment and days between assessments, and author effects. There is also a percentage of variance explained by the model which also controls for the number of independent variables or model complexity. The results are discussed by domain in the following section.

[8] This is called R-squared. An adjusted R-squared is also provided which accounts for the number of variables in the model. A better model explains more variance with fewer independent variables. Given that there are many categorical variables, including multiple authors, the adjusted R-squared is a significant reduction.

[9] Detailed tables of individual regression model results are available on request to ian.edelstein@dcyf.wa.gov

Table 25. Significant effects from multivariate analyses (note: *= p<.05 | dash symbol= no significant effect to report)

Dependent variable (note: each row is a separate model)	Gender effects	Racial group effects	Gender/Race interaction	Age/days between assessments	Institution effects	Author effects (# of authors)	Adjusted Variance explained by model
Att.-Behavior time1-2	-	Black*	Female/Native*	-	GHS*, CF/CTS*, other*	8 of 17*	12%
Att.-Behavior time2-3	-	-	-	-	-	5 of 11*	2%
Aggression time1-2	-	-	-	Days betw.*	GHS*	7/17*	16%
Aggression time2-3	-	-	-	-	-	3/11*	0%
Skills time1-2	-	-	-	Age*	GHS*, CF/CTS*, other*	6/17*	8%
Skills time2-3	-	-	-	-	-	2/11	2%
School time1-2	-	-	Female/Asian*	-	GHS*	6/17*	12%
School time2-3	-	Asian*	-	-	Other*	3/9*	9%
Family time1-2	-	-	-	-	GHS*	6/14*	9%
Family time2-3	-	-	-	-	-	1/7*	7%
Assoc time1-2	-	-	Female/Native* , Female/Black*	Days betw.*	GHS*, other*	8/17*	18%
Assoc time2-3	-	-	-	-	-	4/11*	9%
Alc/Drugs time1-2	Female *	Asian*, Black*	-	Age*	CFs/CTS*	5/17*	16%
Alc/Drugs time2-3	-	-	-	-	-	3/11*	4%
Mental Health time1-2	-	Black*	-	-	-	3/17*	2%
Mental Health time2-3	-	-	-	-	-	1/11*	0%

Discussion

The regression results must be considered in relation to both the overall variance (or lack thereof) within each outcome measure (shown in the boxplots in previous sections), as well as the factor analysis performance.[10] Variables that did not demonstrate adequate factor loadings effectively mean that much of the variance cannot be explained by the underlying factor. The School, Family, Alcohol/Drugs and Mental Health domains all performed poorly in factor analysis.

[10] See: https://dcyf.wa.gov/sites/default/files/pdf/reports/idea-factor-analysis-report.pdf

Generally speaking, a good regression model explains at least 20% of the variance in the outcome. After adjusting for the number of independent variables (the adjusted R-squared), none of the models meet the 20% threshold. The measure of days between assessments could be interpreted as a proxy for the amount of programming a client has received, albeit only between the two assessment periods and without consideration of the actual treatment and services received. Age effects could represent normative, developmental differences in the nature of the domain (e.g. more exposure to alcohol or drugs with age), as well as structural aspects, such as school completion or drop-out, or living independent from family. Gender and racial category effects could represent actual group differences, e.g. cultural differences, but may also indicate bias in either the assessment items, the nature of the responses and their interpretation, and/or potential bias on the part of the assessors.

A well-constructed, well implemented, culturally responsive/appropriate assessment tool should not, overall, exhibit significant subgroup differences. Institution effects, after controlling for age, could suggest differences in programming, as well as the specific effects of the environment or institutional culture. This should be interpreted cautiously as there are structural effects as well: i.e. females and youth under 18 are not GHS residents.

Lastly, author/assessor effects, after controlling for all the other factors, suggest that there are unexplained differences by author. Robust mechanisms for quality assurance (systematically checking the responses and evidence across assessments for accuracy), methods for correction and quality improvement, along with efforts to establish reliability between authors and across time can all work to both reduce or eliminate significant author effects, as well as to improve the accuracy of the data (the signal vs. the noise or error) along with the explanatory power of key independent variables.

Attitude-Behavior Domain

The Attitude-Behavior domain exhibited adequate factor analysis suggesting its potential suitability for change scoring and regression modeling as an outcome. As seen previously, there was significantly more reduction from time 1-2 and for females. In multivariate analysis of time 1-2, there is a significantly higher need in the Black racial group (of +4.7 percentage points above the reference group) and a lower need for female/Native American (-12.5 percentage points).[11] Institutional effects are significant for all categories with respect to the reference, Echo Glen. Additionally, there are substantial author effects, in both higher and lower directions than the reference, with 8 of the 17 first reassessment authors having a significant coefficient, or effect, at the $p<.05$ (95% confidence) level. Seven of those 17 (41%) had an effect at the $p<.01$ (99% confidence) level. The adjusted R-squared of the model is 12%, making it one of the better-explained models across the study although still limited.

In time 2-3, the R-squared for Attitude-Behavior change is only 9% and the only significant factors are author effects for 4 of the 11 authors (two of those are significant at $p<.01$).

Aggression Domain

The Aggression domain also performed adequately in factor analysis. There were significant mean reductions from time 1-2 for both genders, substantially more for females. This was followed by no mean reductions in

[11] Any female interactions should be interpreted cautiously due to small subgroup numbers.

time 2-3. In multivariate testing, the adjusted R-squared for Aggression change from time 1-2 was 16% (among the second highest of the models). There are no significant gender or racial group effects. Days between assessments is predictive of higher aggression needs with an increase of 0.04 percentage points with every additional day ($p<.05$). There is also a significant GHS effect of 27 points higher than Echo Glen, all else being equal. Lastly, there are substantial author effects for 8 of the 17 second assessment authors (7 of those at $p<.01$). There is likely correspondence between the Attitude/Behavior domain and Aggression. Aside from the enduring (and concerning) author effects, the significance of GHS, after controlling for increased age and gendered effects, suggests the possibility that the environment influences increases in aggression, all else being equal.

In time 2-3, the adjusted R-squared for Aggression change is effectively zero and the only significant factors are author effects for 3 of the 11 authors (two of those are significant at $p<.01$).

Social Skills Domain

The Social Skills domain was found to be adequate in factor analysis. A significant mean change reduction was seen from time 1-2 only, with significantly more reduction among females. The adjusted R-squared for Skills change from time 1-2 was only 8%, suggesting relatively poor modeling of the outcome, overall. There were no significant gender or racial group effects. Age at first assessment was significant at $p<.05$ with each increased year predicting 1 percentage point higher Skill need (not Skill acquisition, but in the negative direction). There are significant institutional effects, all higher/greater need than the Echo Glen reference category. This is particularly high for the Other Institution category where location at time of assessment indicated county or state jurisdiction (not JR). There are only 4 such cases for the Skill change time 1-2 variable but the coefficient is a need increase of 52 percentage points, suggesting extreme outliers and potentially a lack of oversight/quality assurance. There are substantial author effects for 6 of the 17 second assessment authors (all at $p<.01$). Collectively, these effects suggest there may be a lack of common understanding of how to assess and score the Skills domain accurately and consistently across time, location and authors.

In time 2-3, the adjusted R-squared for Skills change is 2% and the only significant factors are author effects for 2 of the 11 authors (one of those significant at $p<.01$).

School Domain

The School need domain performed poorly in factor analysis and there were evident structural issues: older clients were often no longer in schooling (and no secondary schooling available at GHS). Meanwhile for Echo Glen residents, school participation was effectively mandatory meaning some School domain measures would automatically improve under incarceration. Thus, it is uncertain if School domain changes can be meaningfully analyzed and interpreted. There were significant mean reductions in School need from time 1-2 for both males and females with no other significant mean differences.

The regression model for School Domain change from time 1-2 had an adjusted R-squared of 12%, a moderate amount across this analysis. The female/Asian category was the only one to exhibit a significant effect (21 percentage points less than the reference, $p<.01$), though this must be interpreted cautiously due to small group sizes. There is a significant need increase for GHS of 19 points ($p<.01$), but this must be considered in light of both the structural issues (e.g. higher age of GHS vs. EG residents, lack of access to schooling) and the

frequency of Not Applicable responses that would render a client's change score impossible to calculate and compare. There are substantial author effects for 6 of the 14 second assessment authors (4 of those at p<.01).

In time 2-3, the adjusted R-squared for School change is 9% (not good, but among the highest for any time 2-3 model). There is a significant Asian racial group effect, an increased School need of 19 percentage points vs. the reference category (again cautiously interpret due to small group size), alongside a significant location effect for only the limited Other category (a reduction of 20 points). Lastly, there were significant effects for 3 of the 9 authors (two of those significant at p<.01).

While participation in schooling, alongside education achievement, are undoubtedly positive outcomes and protective factors against future antisocial behaviors, measurement and subsequent analysis of the School domain does not appear to provide useful or reliable information. It is possible that other measures of educational attachment and achievement could be more accurate and appropriate (but would still be subject to the challenges of consistent implementation fidelity).

Associations Domain

Although there are theoretical reasons to question the viability of measuring change in the Associations domain while incarcerated (when choices of association are inherently limited), this domain performed well under static (one point in time) factor analysis. There was again a significant mean reduction in Associations need from time 1-2 with greater reduction among females.

In regression modeling of Associations domain change from time 1-2, the adjusted R-squared was 18%, among the best fitting models across this analysis. There were significant differences for female/Native American and Female/Black at the p<.05 level relative to the reference group, though these group sizes are small. Female/Native American saw a 16 point greater need reduction compared to the reference and Female/Black, an 11 point greater reduction. Days between assessments had a reduction effect of 0.04 points for every additional day (p<.05). There were institutional effects for GHS (increased need of 7 points, p<.01) and the other location (increase of 34 points, p<.01). Author effects were significant for 8 of the 17 authors (7 of those at p<.01).

In time 2-3, the adjusted R-squared for Associations change is 9%, again among the highest for time 2-3. However, the only significant factors are author effects for 4 of the 11 authors (two of those significant at p<.01). The concerning author and Other institution/location effects notwithstanding, the relative stability of both static factor analysis and the change scoring and regression may suggest that there is a relatively valid construct underlying these Association measures. While group numbers are too small to draw meaningful conclusions from, the apparent improvement for Native American and Black females hints at the possibility of at least a short-term protective effect of institutionalization and removal from negative associations. That is not to suggest that institutionalization is necessarily beneficial for any group or individual.

Family Domain

The Family domain contains many items and exhibited poor factor analysis results. As with the School and Associations domains, there are reasons that Family domain change would not be valid when moving into custody and away from the family environment. That said, there were mean reductions in Family need from time 1-2 across both genders.

In multivariate testing, the Family change from time 1-2 model had an adjusted R-squared of 9%. There were no significant gender or race effects, an increased need among GHS residents (+10%, p<.05), and significant author effects for 6 of the 14 authors (5 of those at p<.01). In time 2-3, the adjusted R-squared for Family change is 7%, with the only significant factor a single author effects of -13 points (p<.01).

As with the School domain, Family aspects are a significant part of the story of each young person's development and acculturation. However, the performance of the Family domain in factor analysis and the change analysis here suggests very little substance, in terms of change, is being consistently captured.

Alcohol and Drug Domain

Alcohol and drug use is undeniably connected with criminal behavior, with potential bidirectional effects. For young, developing bodies and minds, use and abuse of these substances can be extremely damaging and impactful. Studies suggest that a very high percentage of JR clients have alcohol and/or substance use disorders and may require intensive treatment.[12] The IDEA's Alcohol/Drugs Domain is intended to capture these needs with both a historical and current component. However, factor analysis results were poor, and the broad range of items are likely not reflective of a single, underlying clinical condition or need. There is also an obvious structural aspect in that newly incarcerated youth may have their access to alcohol and drugs immediately interrupted-and reflect in decreased current usage.

Indeed, there was a significant mean reduction in Alcohol/Drug Domain needs from time 1-2 for all, significantly more reduction for females (this could also reflect a difference in access between EG and GHS). The regression model for Alcohol/Drug Domain change from time 1-2 had an adjusted R-squared of 16%. There was a significant gender effect of -7 percentage points for females (p=.05), all else being equal. There are also racial group effects with Asians 6.5 points less in need and Black 6.7 points higher in need (both p<.05). As anticipated, age has a significant effect at +2.5 points for each additional year of age (p<.01). Coefficients for all the institutions/location settings beyond EG are indicative of increased Alcohol/Drug need, although only the CF/CTS category is significant (+6 points, p<.05). There are author effects for 5 of the 17 authors (4 at p<.01).

In time 2-3, the adjusted R-squared for Alcohol-Drug domain change is 4%, with author effects the only significant independent variables (3, all at p<.01).

Overall, the racial group differences do create some questions or concern, particularly the significant and opposing directions between Asian and Black groups. Group numbers are sufficient for interpretation although the Black group (n=190 for this variable) is substantially larger than the Asian (n=32). This domain is one that may benefit from a more stringent clinical assessment, or at least the use of a clinical/methodological approach to ensure that scores are as accurate and consistent as possible and do not suffer from subjectivity or bias.

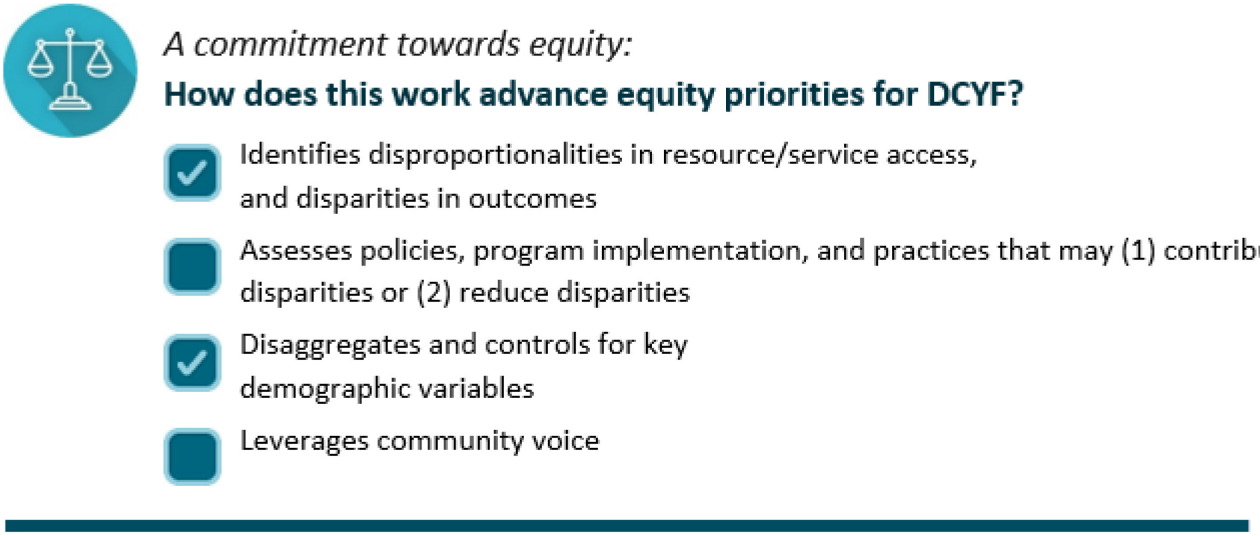
Mental Health Domain

Finally, the Mental Health domain saw the worst factor analysis results and could not be considered reliable for static evaluation, let alone dynamic change scoring. There were no significant mean reductions or

[12] Roughly 75% of all JR clients indicate medium to high SUD treatment needs at intake.

increases in the prior descriptive analyses. In multivariate modeling, the adjusted R-squared for Mental Health domain change from time 1-2 was only 2%, the worst among all time 1-2 models and zero for time 2-3. In time 1-2, there was a significant Mental Health need increase for Black persons of 4 percentage points (p<.01), in addition to 3 author effects (2 at P<.01). In time 2-3, the only significant effect was one author at p<.01.

The range of items included in the Mental Health domain (even after reducing items to improve factor analysis reliability somewhat) do not appear, at face or theoretical levels, to cohere to specific issues or diagnoses. There are an overwhelming number of historical (ever experienced or witnessed) items that are not dynamic in nature, nor do they consider how a client feels or thinks vs. adverse events they've been through. Thus, the IDEA Mental Health domain should not be considered a reliable source of information nor should any change scoring. That said, the significance of the reduction in need only for Black clients may hint at either a biased perception of reduced mental health issues or needs, or the desire for Black clients to underreport their needs.



Learn more about these areas of focus