Disruptive Behavior Disorder Meta-analysis

November 15, 2014

```
In [1]: %matplotlib inline
        import matplotlib.pyplot as plt
        import seaborn as sb
        import numpy as np
        import pandas as pd
        import pymc as pm
        from scipy.misc import comb
```

```
In [2]: np.random.seed(42)
```

0.1 Data import and cleaning

```
In [3]: study_char = pd.read_excel("DBD Data for Meta Analyses.xlsx", "Study Characteristics",
                                   index_col='RefID', na_values=['-', 'NR'])
        outcomes = pd.read_excel("DBD Data for Meta Analyses.xlsx", "Outcomes",
                                   na_values=['ND', 'NR'])
        demographics = pd.read_excel("DBD Data for Meta Analyses.xlsx", "Pt Demographics", na_values=['-
```

Data cleaning

```
In [4]: # Cast outcomes variables to floats
        for col in ('Last FU Mean', 'Last FU SD',):
            outcomes[col] = outcomes[col].astype(float)
```

```
In [5]: # Recode age category
        study_char['age_cat'] = study_char.AgeCat.replace({'PRE-K':1, 'SCHOOL':0, 'TEEN':2})
```

```
In [6]: # Fix data label typo
        outcomes['Measure Instrument'] = outcomes['Measure Instrument'].replace({'Eyberg Child Behaviou
                                            'Eyberg Child Behaviour Inventory, Problem Subscale'})
        outcomes.Units = outcomes.Units.replace({'scale': 'Scale'})
```

```
In [7]: # Parse followup times and convert to months
        split_fut = outcomes.loc[outcomes['Last FU Time'].notnull(), 'Last FU Time'].apply(lambda x: st
        fut_months = [float(time)/52.*(unit=='weeks') or float(time) for time, unit in split_fut]
        outcomes.loc[outcomes['Last FU Time'].notnull(), 'Last FU Time'] = fut_months
```

We are assumung all CBC Externalizing values over 50 are T-scores, and those under 50 are raw scores. This recodes those observations.

```
In [8]: cbce_ind = outcomes['Measure Instrument'].apply(lambda x: x.startswith('Child Behavior Checklis
        under_50 = outcomes['BL Mean']<50
        outcomes.loc[cbce_ind & (under_50^True), 'Measure Instrument'] = 'Child Behavior Checklist, Exte
        outcomes.loc[cbce_ind & under_50, 'Measure Instrument'] = 'Child Behavior Checklist, Externaliz
```

Recode measure instrument variables

```
In [9]: instrument = []
        subtype = []
        units = []

        for i,row in outcomes.iterrows():
            separator = row['Measure Instrument'].find(',')
            if separator == -1:
                separator = row['Measure Instrument'].find('-')
            instrument.append(row['Measure Instrument'][:separator])
            s = row['Measure Instrument'][separator+2:]
            paren = s.find('(')
            if paren > -1:
                subtype.append(s[:paren-1])
                units.append(s[paren+1:-1])
            else:
                subtype.append(s)
                if s.endswith('scale'):
                    units.append('Scale')
                else:
                    units.append('Score')

        new_cols = pd.DataFrame({'instrument': instrument, 'subtype': subtype,
                                 'units': units}, index=outcomes.index)
```

```
In [10]: outcomes['Measure Instrument'].value_counts()
```

```
Out[10]: Eyberg Child Behaviour Inventory, Intensity Subscale            63
         Eyberg Child Behaviour Inventory, Problem Subscale              45
         Child Behavior Checklist, Externalizing (T Score)               33
         Child Behavior Checklist, Externalizing                         11
         Eyberg Child Behaviour Inventory, Intensity Subscale (T Score)    10
         Strengths and Difficulties Questionnaire- Conduct Problems Scale    10
         Child Behavior Checklist, Aggression                             4
         Strengths and Difficulties Questionnaire- Emotional Symptoms Scale     4
         Strengths and Difficulties Questionnaire- Total Difficulties Score     4
         Strengths and Difficulties Questionnaire- Total Score            4
         Eyberg Child Behaviour Inventory, Problem Subscale (T Score)      4
         Strengths and Difficulties Questionnaire- Impact Score           2
         Strengths and Difficulties Questionnaire- Hyperactivity Scale     2
         Child Behavior Checklist, Conduct Problems                       2
         Child Behavior Checklist, Rulebreaking                           2
         Child Behavior Checklist, Conduct Problems (T Score)             2
         dtype: int64
```

```
In [11]: new_cols.head()
```

```
Out[11]:                         instrument              subtype    units
         0  Eyberg Child Behaviour Inventory  Intensity Subscale  T Score
         1  Eyberg Child Behaviour Inventory    Problem Subscale  T Score
         2  Eyberg Child Behaviour Inventory  Intensity Subscale  T Score
         3  Eyberg Child Behaviour Inventory    Problem Subscale  T Score
         4          Child Behavior Checklist       Externalizing    Score
```

```
In [12]: # Append new columns
         outcomes = outcomes.join(new_cols)
```

```
In [13]: outcomes.intvn.value_counts()
```

```
Out[13]: wlc              41
         tau              37
         iypt             26
         pcit             16
         pppsd             7
         iyptndiyct        6
         mst               5
         it                4
         ppcp              4
         pppo              4
         pmto              3
         pcitc             3
         snap              3
         spokes            3
         pmtndp            3
         pppe              3
         pmtpa             2
         iyct              2
         hncte             2
         setpc             2
         pcitabb           2
         modularndn        2
         pmtsd             2
         hncstd            2
         pmtnds            2
         cpp               1
         cbt               1
         pppstd            1
         coaching          1
         mcfi              1
         modularndcomm     1
         itpt              1
         kitkashrut        1
         sst               1
         pstnds            1
         hnc               1
         hitkashrut        1
         iyptadv           1
         scip              1
         modularndclin     1
         projndsupport     1
         dtype: int64
```

0.2 Data summaries

Cross-tabulation of the outcome counts by measure instrument

```
In [14]: pd.crosstab(outcomes['instrument'], outcomes['Outcome'])
```

```
Out[14]: Outcome                                    01 Behavior, disruptive  \
         instrument
         Child Behavior Checklist                                       48
         Eyberg Child Behaviour Inventory                              122
```

```
Strengths and Difficulties Questionnaire                      10

Outcome                                 02 Behavior, aggression  \
instrument
Child Behavior Checklist                                       4
Eyberg Child Behaviour Inventory                               0
Strengths and Difficulties Questionnaire                       0

Outcome                                 06 Behavior, fighting, destruction, violation  \
instrument
Child Behavior Checklist                                                             2
Eyberg Child Behaviour Inventory                                                     0
Strengths and Difficulties Questionnaire                                             0

Outcome                                 08 Behavior, other
instrument
Child Behavior Checklist                                 0
Eyberg Child Behaviour Inventory                         0
Strengths and Difficulties Questionnaire                16
```

Distribution of age categories

In [15]:
```
study_char.AgeCat.value_counts()
```

Out[15]:
```
SCHOOL    46
PRE-K     26
TEEN      14
dtype: int64
```

Frequencies of various intervention types

In [16]:
```
study_char['Intervention Type'].value_counts()
```

Out[16]:
```
PHARM                                                          20
IY-PT                                                           8
MST                                                             7
PCIT                                                            5
IY-PT + IY-CT                                                   4
BSFT                                                            3
PMTO                                                            3
Triple P (enhanced)                                             3
PT                                                              2
Fast Track                                                      2
PCIT-ABB                                                        2
Triple-P (self-directed)                                        2
CBT                                                             2
IY-CT                                                           2
IY-PT (nurse led)                                               2
OTH: Intensive treatment                                        2
PMT (practitioner assisted)                                     1
HNC                                                             1
OTH: Modular (nurse administered)                               1
MF-PEP + TAU                                                    1
SNAP Under 12 OutReach Project (enhanced)                       1
OTH: Booster                                                    1
```

```
Coping power                                                     1
OTH: Child only treatment                                        1
OTH: Family therapy                                              1
SCIP (Social cognitive (Dodge's))                                1
Triple-P (enhanced)                                              1
PONI                                                             1
OTH: Parental Stress                                             1
PCIT (modified)                                                  1
Triple-P (online)                                                1
UCPP                                                             1
IY-PT (brief)                                                    1
OTH: Modular treatment (community)                               1
MST (PIT)                                                        1
OTH: Modular treatment                                           1
SNAP Under 12 \nOutReach Project(ORP)                            1
OTH: Instrumental, emotional support & child management skills    1
IY-PT + IY-CT + IY-TT                                            1
PMT (perceptive)                                                 1
PHARM1 + PHARM2                                                  1
HNC (technology enhanced)                                        1
OTH: Community Parent Education Program                          1
Coping Power                                                     1
Coping power; Coping Power + Booser                              1
IY-CT + IY-PT                                                    1
OTH: FFT                                                         1
OTH: Day Program                                                 1
OTH: Parents Plus Children's Program                             1
IY-PT + ADVANCE                                                  1
OTH: Project support                                             1
CPS                                                              1
Coping Power (cultural adaptation)                               1
SNAP Under 12 OutReach Project                                   1
Parenting Group (SPOKES)                                         1
SET-PC                                                           1
PHARM + PSYCH                                                    1
Length: 57, dtype: int64
```

0.3 Extract variables of interest and merge tables

In [17]:
```
KQ1 = study_char[study_char.KQ=='KQ1']
```

In [18]:
```
study_varnames = ['Year', 'age_cat', 'Geographic setting', 'Age mean (years) ', 'Age SD (years)
              'Age min (years)', 'Age max (years)', 'Proportion Male (%)']

study_vars = KQ1[study_varnames].rename(columns={'Geographic setting': 'country',
                             'Age mean (years) ': 'age_mean',
                             'Age SD (years)': 'age_sd',
                             'Age min (years)': 'age_min',
                             'Age max (years)': 'age_max',
                             'Proportion Male (%)': 'p_male'})
```

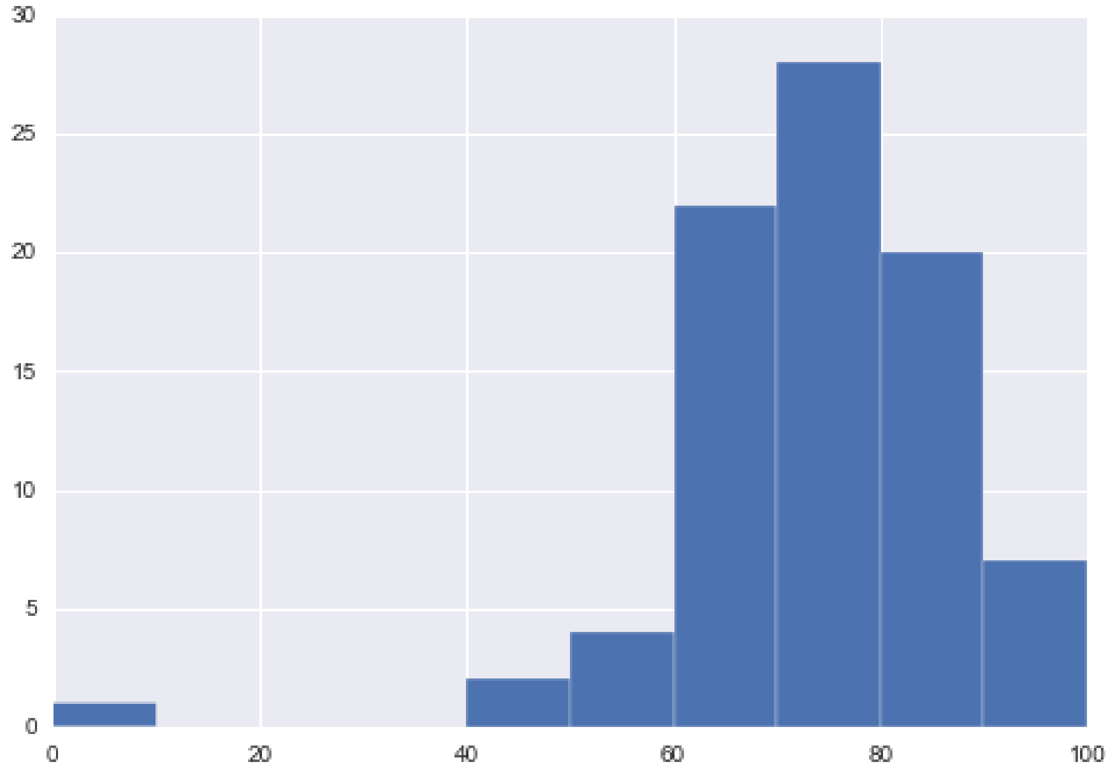
In [19]:
```
study_vars.head()
```

Out[19]:
```
        Year  age_cat     country  age_mean  age_sd  age_min  age_max  p_male
RefID
```

```
23   2013        1          USA   2.80   0.61        2        4      62
100  2013        2          USA  14.60   1.30       11       18      83
103  2013        1          USA   5.67   1.72        3        8      53
141  2013        0          USA   9.90   1.30        8       11      73
156  2013        2  Netherlands  16.00   1.31       12       18      73
```

In [20]:
```
study_vars.p_male.hist()
```

Out[20]: <matplotlib.axes._subplots.AxesSubplot at 0x1088dcc50>



Proportion missing

In [21]:
```
study_vars.isnull().mean(0).round(2)
```

Out[21]:
```
Year        0.00
age_cat     0.00
country     0.00
age_mean    0.13
age_sd      0.20
age_min     0.07
age_max     0.08
p_male      0.01
dtype: float64
```

Will assume the mean age for those which are missing is simply the midpoint between minimum and maximum values

```
In [22]: est_means = study_vars.apply(lambda x: x.age_min + (x.age_max - x.age_min) / 2, axis=1)[study_v
         study_vars.loc[study_vars.age_mean.isnull(), 'age_mean'] = est_means

         study_vars.age_mean.isnull().sum()
```

```
Out[22]: 3
```

```
In [23]: outcomes_varnames = ['Ref ID', 'Measure Instrument', 'instrument', 'subtype', 'units',
                              'intvn', 'cc', 'pc', 'fc',
                              'BL N', 'BL Mean', 'BL SD',
                              'EOT \nN', 'EOT Mean', 'EOT \nSD', 'Last FU Time', 'Last FU N',
                              'Last FU Mean', 'Last FU SD', 'CS Group N', 'CS Mean', 'CS SD']
```

```
In [24]: outcomes_vars = outcomes[outcomes_varnames].rename(columns={'Ref ID': 'RefID',
                                                                     'Measure Instrument': 'm
                                                                     'cc': 'child_component'
                                                                     'pc': 'parent_component
                                                                     'fc': 'family_component
                                                                     'oc': 'other_component'
                                                                     'BL N': 'baseline_n',
                                                                     'BL Mean': 'baseline_mea
                                                                     'BL SD': 'baseline_sd',
                                                                     'EOT \nN': 'end_treat_n
                                                                     'EOT Mean': 'end_treat_m
                                                                     'EOT \nSD': 'end_treat_s
                                                                     'Last FU Time': 'followu
                                                                     'Last FU N': 'followup_n
                                                                     'Last FU Mean': 'followu
                                                                     'Last FU SD': 'followup_
                                                                     'CS Group N': 'change_n
                                                                     'CS Mean': 'change_mean
                                                                     'CS SD': 'change_sd'})
```

Recode intervention clasification

```
In [25]: control = ((outcomes_vars.child_component^True) &
                    (outcomes_vars.parent_component^True) &
                    (outcomes_vars.family_component^True)).astype(int)
         child_only = ((outcomes_vars.child_component) &
                       (outcomes_vars.parent_component^True) &
                       (outcomes_vars.family_component^True)).astype(int)
         parent_only = ((outcomes_vars.child_component^True) &
                        (outcomes_vars.parent_component) &
                        (outcomes_vars.family_component^True)).astype(int)
         outcomes_vars.ix[child_only.astype(bool), ['child_component', 'parent_component', 'family_compo
```

Out[25]:

	child_component	parent_component	family_component
112	1	0	0
113	1	0	0
115	1	0	0
116	1	0	0
149	1	0	0
173	1	0	0

```
In [26]: multi_component = ((parent_only^True) & (child_only^True) & (control^True)).astype(int)
```

```
outcomes_vars['child_only'] = child_only
outcomes_vars['parent_only'] = parent_only
outcomes_vars['multi_component'] = multi_component
```

Obtain subset with non-missing EOT data

```
In [27]: eot_subset = outcomes_vars[outcomes_vars.end_treat_mean.notnull() & outcomes_vars.end_treat_sd
```

Calculate EOT difference

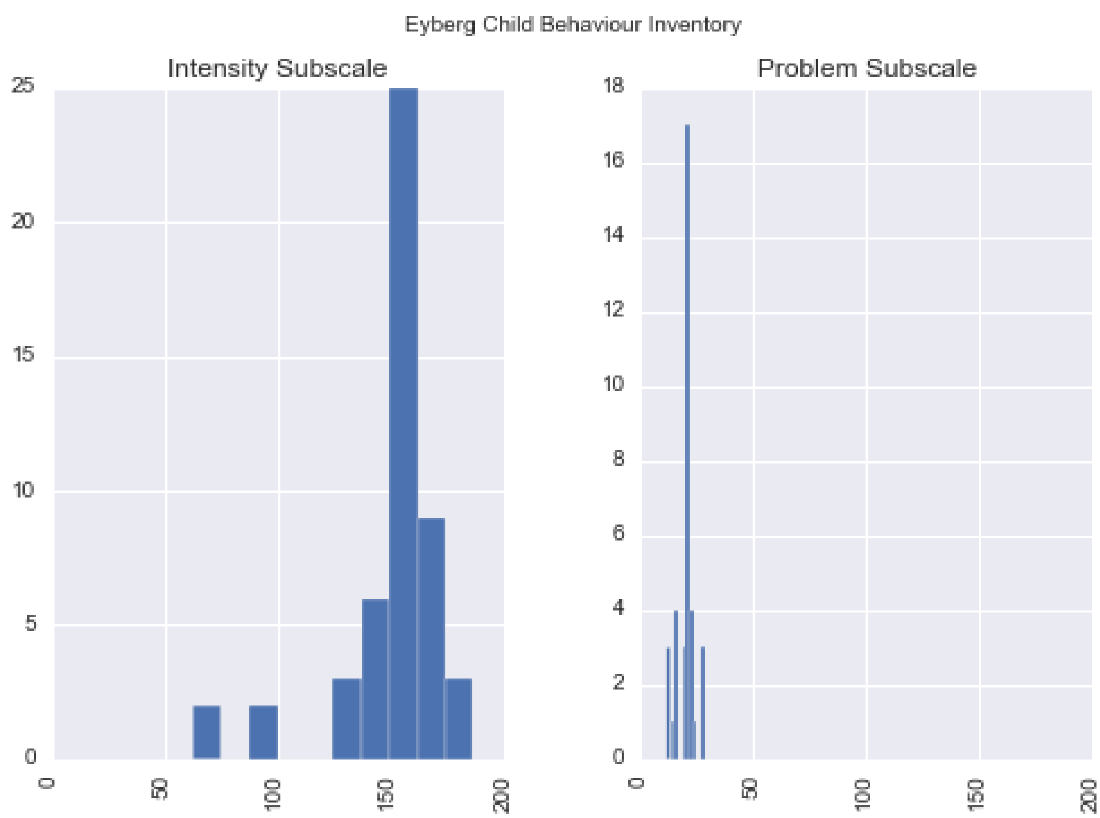
```
In [28]: eot_subset['eot_diff_mean'] = eot_subset.end_treat_mean - eot_subset.baseline_mean
```

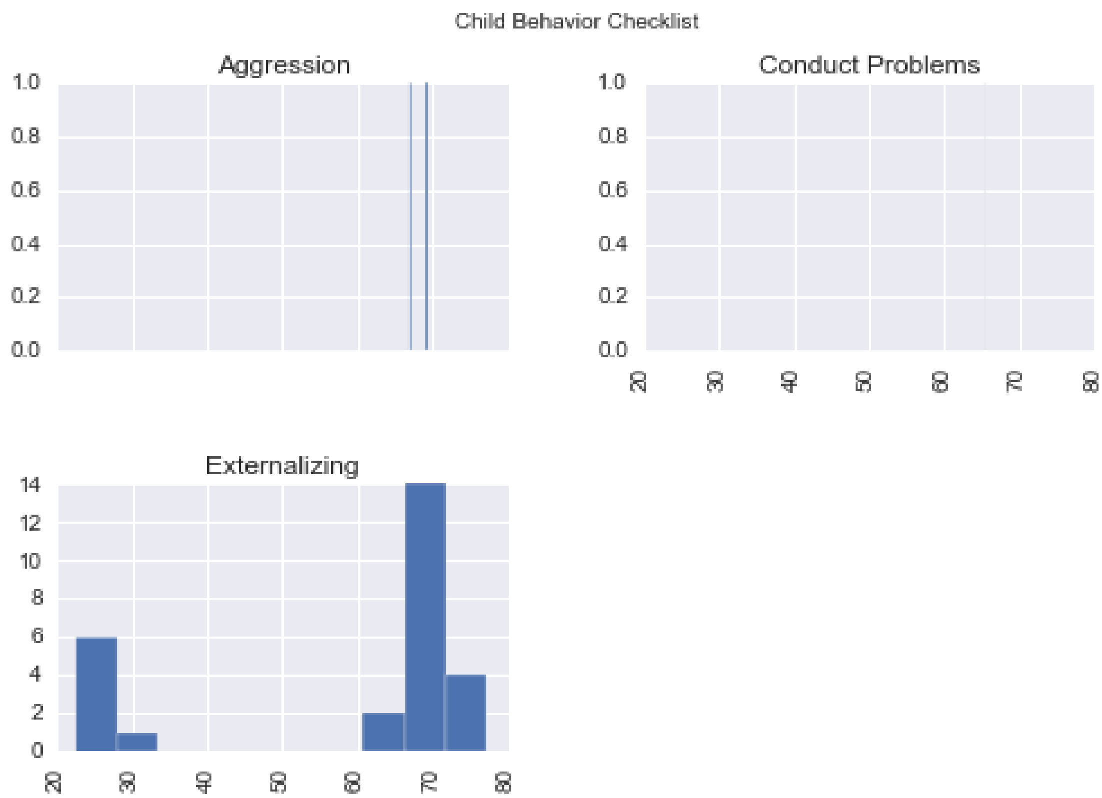
```
/usr/local/lib/python3.4/site-packages/IPython/kernel/__main__.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#inde
  if __name__ == '__main__':
```

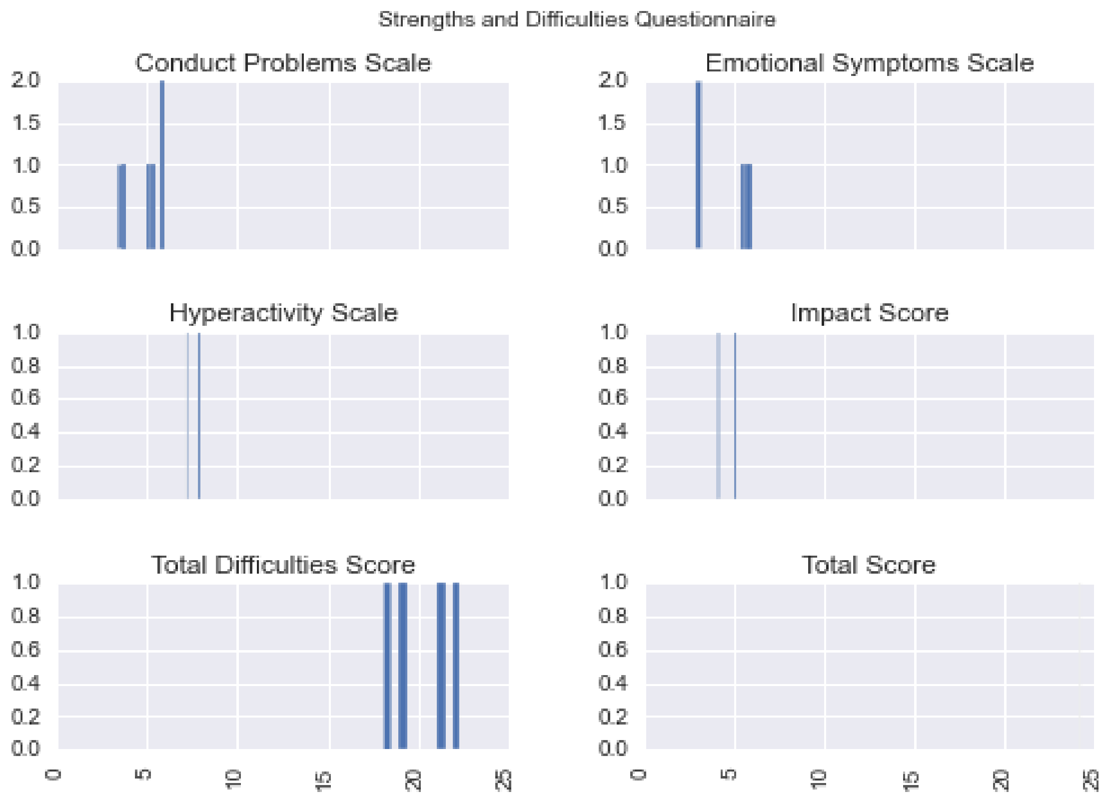
```
In [29]: eot_subset['eot_diff_sd'] = eot_subset.baseline_sd + eot_subset.end_treat_sd

In [30]: eot_subset['eot_diff_n'] = eot_subset[['baseline_n', 'end_treat_n']].min(1)
```

Distribution of baseline means among outcome metrics

```
In [31]: for instrument in ('Eyberg Child Behaviour Inventory',
                           'Child Behavior Checklist',
                           'Strengths and Difficulties Questionnaire'):
             eot_subset[eot_subset.instrument==instrument]['baseline_mean'].hist(by=eot_subset['subtype
                                                                                  sharex=True)
             plt.suptitle(instrument);
```




Child Behavior Checklist
Aggression
1.0
0.8
0.6
0.4
0.2
0.0
Conduct Problems
1.0
0.8
0.6
0.4
0.2
0.0
20
30
40
50
60
70
80
Externalizing
14
12
10
8
6
4
2
0
20
30
40
50
60
70
80



In [32]: `eot_subset.instrument.value_counts()`

Out[32]:
```
Eyberg Child Behaviour Inventory            86
Child Behavior Checklist                    31
Strengths and Difficulties Questionnaire    20
dtype: int64
```

In [33]: `eot_subset[eot_subset.RefID==441]`

Out[33]:
```
    RefID                                   measure_instrument  \
14    441  Eyberg Child Behaviour Inventory, Intensity Su...
15    441  Eyberg Child Behaviour Inventory, Problem Subs...
16    441  Eyberg Child Behaviour Inventory, Intensity Su...
17    441  Eyberg Child Behaviour Inventory, Problem Subs...

                          instrument              subtype  units intvn  \
14  Eyberg Child Behaviour Inventory  Intensity Subscale  Scale  iypt
15  Eyberg Child Behaviour Inventory    Problem Subscale  Scale  iypt
16  Eyberg Child Behaviour Inventory  Intensity Subscale  Scale   wlc
17  Eyberg Child Behaviour Inventory    Problem Subscale  Scale   wlc

    child_component  parent_component  family_component  baseline_n  \
14                0                 1                 0          32
15                0                 1                 0          24
16                0                 0                 0          20
17                0                 0                 0          17
```

```
    ...     followup_sd  change_n  change_mean  change_sd  child_only  \
14  ...             NaN       NaN          NaN        NaN           0
15  ...             NaN       NaN          NaN        NaN           0
16  ...             NaN       NaN          NaN        NaN           0
17  ...             NaN       NaN          NaN        NaN           0

   parent_only  multi_component  eot_diff_mean  eot_diff_sd eot_diff_n
14           1                0         -31.40        46.80         32
15           1                0          -9.70        12.02         24
16           0                0          -5.80        49.60         20
17           0                0          -2.88        14.59         17

[4 rows x 28 columns]
```

Several studies use multiple instruments and metrics within instruments

```
In [34]: eot_subset.groupby(['RefID', 'instrument'])['subtype'].value_counts()
```

```
Out[34]: RefID  instrument
         441    Eyberg Child Behaviour Inventory          Problem Subscale           2
                                                          Intensity Subscale         2
         475    Eyberg Child Behaviour Inventory          Intensity Subscale         2
         539    Child Behavior Checklist                  Externalizing              2
                                                          Aggression                 2
         564    Eyberg Child Behaviour Inventory          Problem Subscale           2
                                                          Intensity Subscale         2
         899    Child Behavior Checklist                  Externalizing              3
                Eyberg Child Behaviour Inventory          Problem Subscale           3
                                                          Intensity Subscale         3
         993    Strengths and Difficulties Questionnaire  Total Difficulties Score   2
                                                          Impact Score               2
                                                          Conduct Problems Scale     2
                                                          Hyperactivity Scale        2
         1236   Eyberg Child Behaviour Inventory          Problem Subscale           6
                                                          Intensity Subscale         6
         1245   Child Behavior Checklist                  Externalizing              2
                Eyberg Child Behaviour Inventory          Intensity Subscale         2
         1511   Child Behavior Checklist                  Externalizing              2
         1585   Eyberg Child Behaviour Inventory          Problem Subscale           2
                                                          Intensity Subscale         2
         1875   Child Behavior Checklist                  Externalizing              3
         1951   Child Behavior Checklist                  Externalizing              2
         2092   Child Behavior Checklist                  Externalizing              3
                Eyberg Child Behaviour Inventory          Intensity Subscale         3
         2117   Child Behavior Checklist                  Externalizing              2
         2219   Eyberg Child Behaviour Inventory          Problem Subscale           3
                                                          Intensity Subscale         3
         2239   Child Behavior Checklist                  Externalizing              2
         2347   Strengths and Difficulties Questionnaire  Total Score                2
                                                          Emotional Symptoms Scale   2
                                                          Conduct Problems Scale     2
         3211   Eyberg Child Behaviour Inventory          Problem Subscale           2
                                                          Intensity Subscale         2
                Strengths and Difficulties Questionnaire  Total Difficulties Score   2
```

```
3225   Eyberg Child Behaviour Inventory         Problem Subscale           2
                                                Intensity Subscale         2
3397   Child Behavior Checklist                 Externalizing              2
3399   Child Behavior Checklist                 Externalizing              2
3495   Eyberg Child Behaviour Inventory         Intensity Subscale         4
3687   Eyberg Child Behaviour Inventory         Problem Subscale           6
                                                Intensity Subscale         6
3716   Eyberg Child Behaviour Inventory         Problem Subscale           2
                                                Intensity Subscale         2
3766   Child Behavior Checklist                 Conduct Problems           2
       Eyberg Child Behaviour Inventory         Intensity Subscale         5
3915   Eyberg Child Behaviour Inventory         Problem Subscale           2
                                                Intensity Subscale         2
3960   Eyberg Child Behaviour Inventory         Problem Subscale           2
7109   Eyberg Child Behaviour Inventory         Problem Subscale           2
                                                Intensity Subscale         2
       Strengths and Difficulties Questionnaire Emotional Symptoms Scale   2
                                                Conduct Problems Scale     2
7723   Child Behavior Checklist                 Externalizing              2
Length: 54, dtype: int64
```

In [35]:
```
pd.crosstab(eot_subset.instrument, eot_subset.subtype)
```

Out[35]:
```
subtype                                   Aggression  Conduct Problems  \
instrument
Child Behavior Checklist                           2                 2
Eyberg Child Behaviour Inventory                   0                 0
Strengths and Difficulties Questionnaire           0                 0

subtype                                   Conduct Problems Scale  \
instrument
Child Behavior Checklist                                       0
Eyberg Child Behaviour Inventory                               0
Strengths and Difficulties Questionnaire                       6

subtype                                   Emotional Symptoms Scale  \
instrument
Child Behavior Checklist                                         0
Eyberg Child Behaviour Inventory                                 0
Strengths and Difficulties Questionnaire                         4

subtype                                   Externalizing  Hyperactivity Scale  \
instrument
Child Behavior Checklist                             27                    0
Eyberg Child Behaviour Inventory                      0                    0
Strengths and Difficulties Questionnaire              0                    2

subtype                                   Impact Score  Intensity Subscale  \
instrument
Child Behavior Checklist                             0                   0
Eyberg Child Behaviour Inventory                     0                  50
Strengths and Difficulties Questionnaire             2                   0

subtype                                   Problem Subscale  \
instrument
```

```
Child Behavior Checklist                                 0
Eyberg Child Behaviour Inventory                        36
Strengths and Difficulties Questionnaire                 0

subtype                                   Total Difficulties Score  \
instrument
Child Behavior Checklist                                         0
Eyberg Child Behaviour Inventory                                 0
Strengths and Difficulties Questionnaire                         4

subtype                                   Total Score
instrument
Child Behavior Checklist                            0
Eyberg Child Behaviour Inventory                    0
Strengths and Difficulties Questionnaire            2
```

In [36]:
```
x = eot_subset[eot_subset.instrument=='Eyberg Child Behaviour Inventory']
pd.crosstab(x.instrument, x.subtype)
```

Out[36]:
```
subtype                           Intensity Subscale  Problem Subscale
instrument
Eyberg Child Behaviour Inventory                  50                36
```

In [37]:
```
x = eot_subset[eot_subset.instrument=='Child Behavior Checklist']
pd.crosstab(x.instrument, x.subtype)
```

Out[37]:
```
subtype                   Aggression  Conduct Problems  Externalizing
instrument
Child Behavior Checklist           2                 2             27
```

In [38]:
```
x = eot_subset[eot_subset.instrument=='Strengths and Difficulties Questionnaire']
pd.crosstab(x.instrument, x.subtype)
```

Out[38]:
```
subtype                                   Conduct Problems Scale  \
instrument
Strengths and Difficulties Questionnaire                       6

subtype                                   Emotional Symptoms Scale  \
instrument
Strengths and Difficulties Questionnaire                         4

subtype                                   Hyperactivity Scale  Impact Score  \
instrument
Strengths and Difficulties Questionnaire                    2             2

subtype                                   Total Difficulties Score  \
instrument
Strengths and Difficulties Questionnaire                         4

subtype                                   Total Score
instrument
Strengths and Difficulties Questionnaire            2
```

Merge study variables and outcomes

In [39]:
```
merged_vars = study_vars.merge(eot_subset, left_index=True, right_on='RefID')
merged_vars.shape
```

```
Out[39]: (137, 36)
```

For now, restrict to the three most prevalent metrics.

```
In [40]: merged_vars.measure_instrument.value_counts()
```

```
Out[40]: Eyberg Child Behaviour Inventory, Intensity Subscale            46
         Eyberg Child Behaviour Inventory, Problem Subscale              36
         Child Behavior Checklist, Externalizing (T Score)               20
         Child Behavior Checklist, Externalizing                          7
         Strengths and Difficulties Questionnaire- Conduct Problems Scale     6
         Strengths and Difficulties Questionnaire- Emotional Symptoms Scale     4
         Strengths and Difficulties Questionnaire- Total Difficulties Score     4
         Eyberg Child Behaviour Inventory, Intensity Subscale (T Score)     4
         Child Behavior Checklist, Aggression                             2
         Strengths and Difficulties Questionnaire- Total Score            2
         Strengths and Difficulties Questionnaire- Hyperactivity Scale     2
         Child Behavior Checklist, Conduct Problems (T Score)             2
         Strengths and Difficulties Questionnaire- Impact Score           2
         dtype: int64
```

```
In [41]: analysis_subset = merged_vars[merged_vars.measure_instrument.isin(merged_vars.measure_instrume
         analysis_subset.groupby('measure_instrument')['baseline_mean'].max()
```

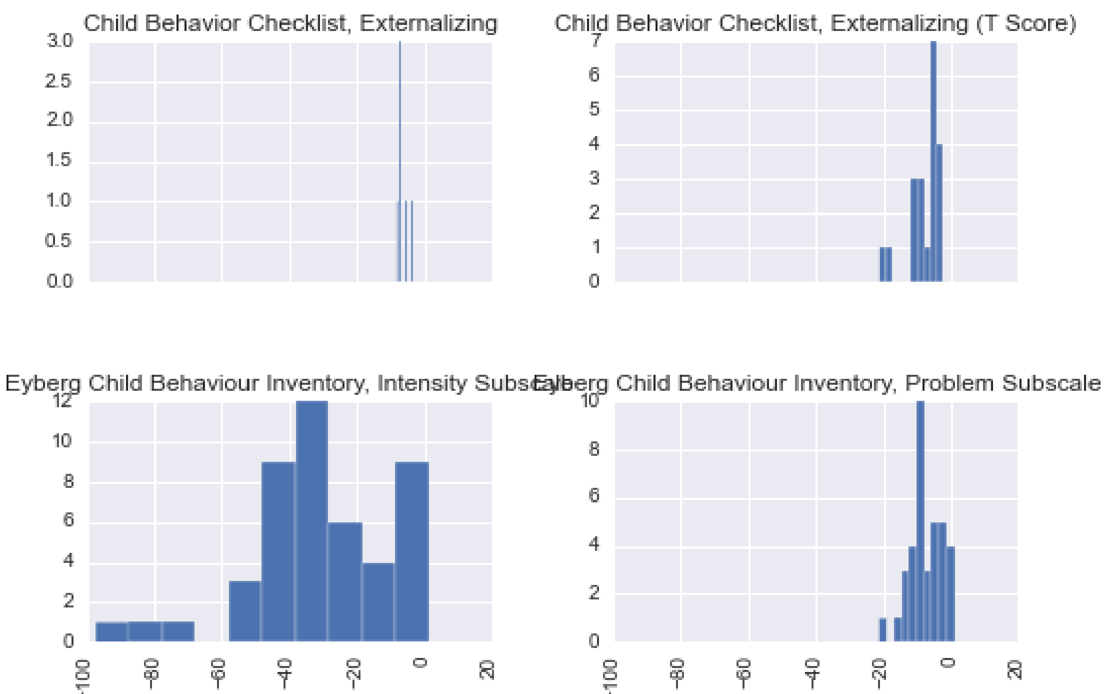
```
Out[41]: measure_instrument
         Child Behavior Checklist, Externalizing                  30.90
         Child Behavior Checklist, Externalizing (T Score)        77.10
         Eyberg Child Behaviour Inventory, Intensity Subscale    186.44
         Eyberg Child Behaviour Inventory, Problem Subscale       28.62
         Name: baseline_mean, dtype: float64
```

```
In [42]: analysis_subset.groupby('measure_instrument').baseline_mean.hist()
```

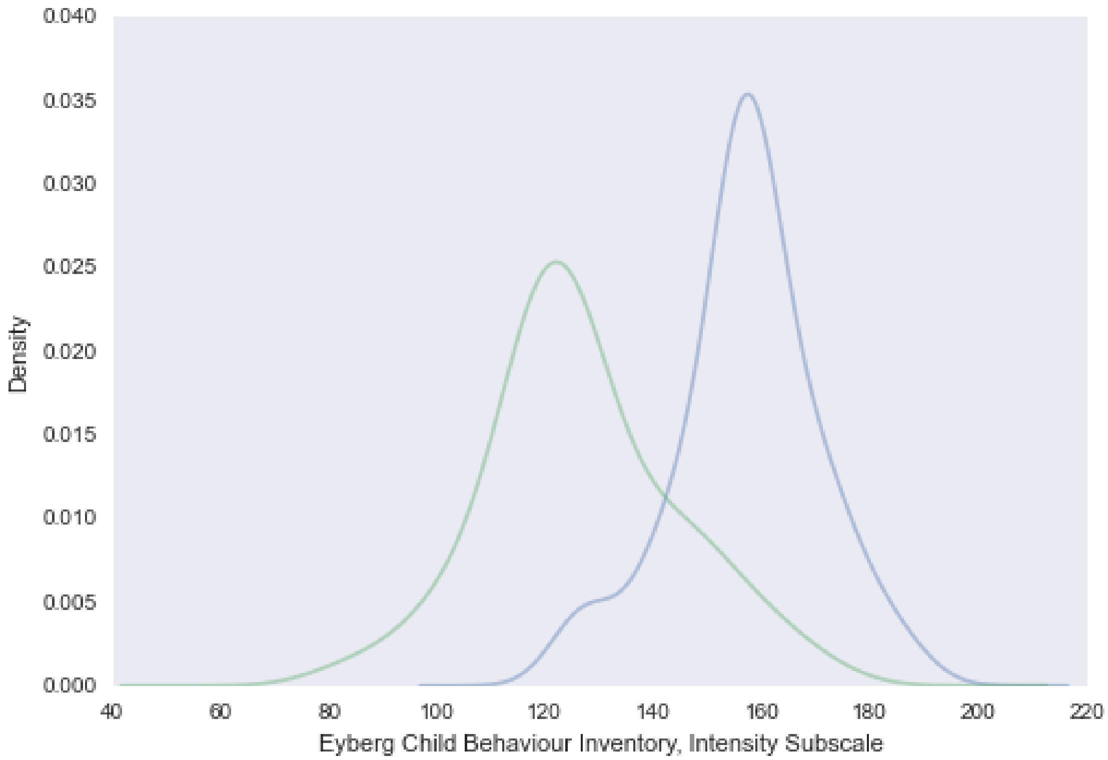
```
Out[42]: measure_instrument
         Child Behavior Checklist, Externalizing                 Axes(0.125,0.125;0.775x0.775)
         Child Behavior Checklist, Externalizing (T Score)       Axes(0.125,0.125;0.775x0.775)
         Eyberg Child Behaviour Inventory, Intensity Subscale    Axes(0.125,0.125;0.775x0.775)
         Eyberg Child Behaviour Inventory, Problem Subscale      Axes(0.125,0.125;0.775x0.775)
         Name: baseline_mean, dtype: object
```

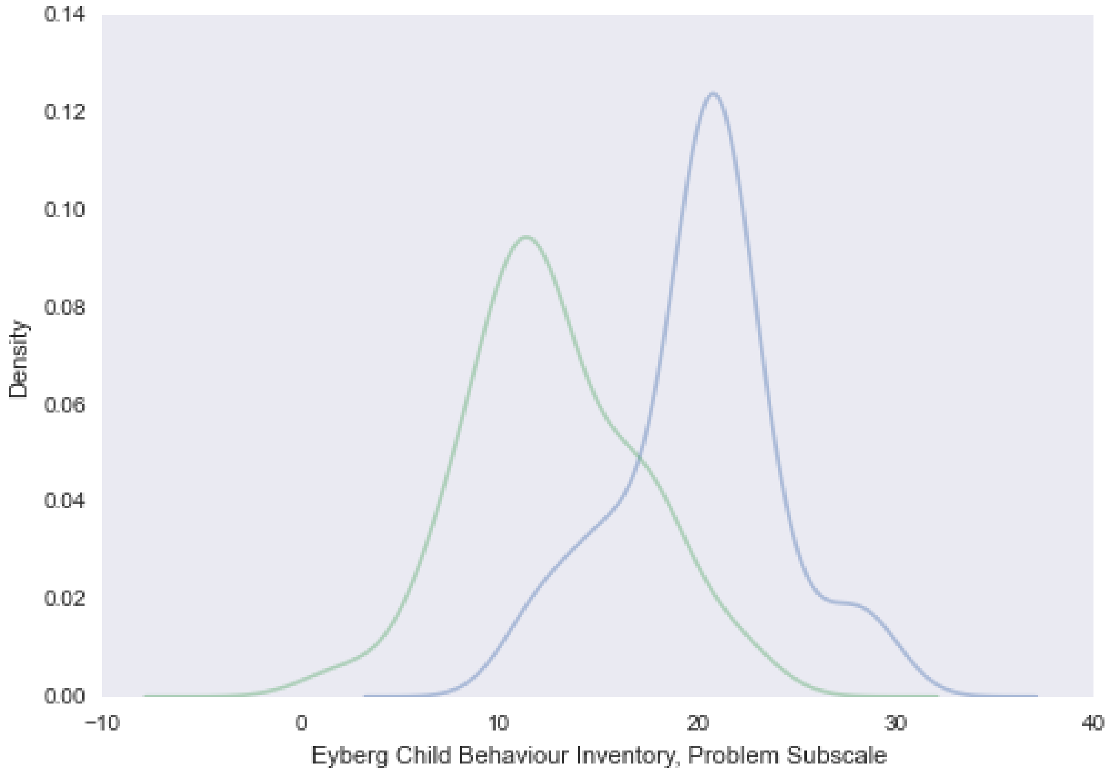

```
In [43]: analysis_subset['baseline_mean'].hist(by=analysis_subset['measure_instrument'],sharex=True)

Out[43]: array([[<matplotlib.axes._subplots.AxesSubplot object at 0x1091950b8>,
                 <matplotlib.axes._subplots.AxesSubplot object at 0x108d18da0>],
                [<matplotlib.axes._subplots.AxesSubplot object at 0x108bab710>,
                 <matplotlib.axes._subplots.AxesSubplot object at 0x109028c18>]], dtype=object)
```

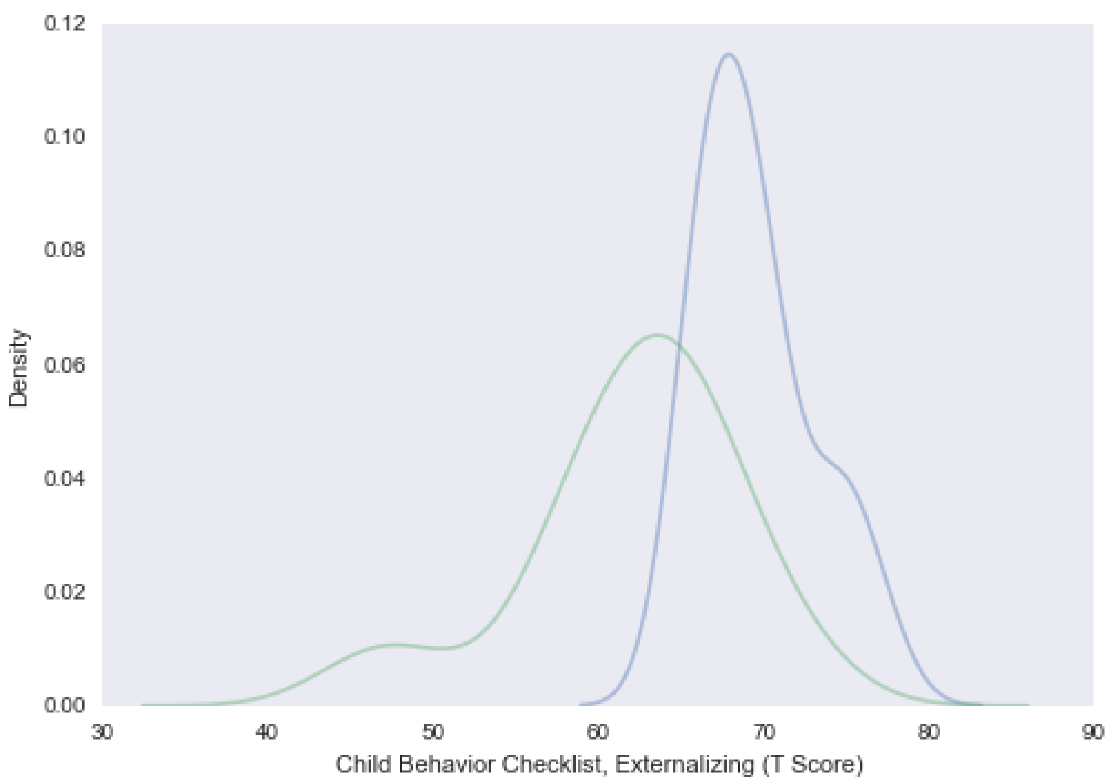

```
In [44]: analysis_subset['eot_diff_mean'].hist(by=analysis_subset['measure_instrument'],sharex=True)

Out[44]: array([[<matplotlib.axes._subplots.AxesSubplot object at 0x108b08a90>,
                 <matplotlib.axes._subplots.AxesSubplot object at 0x108753550>],
                [<matplotlib.axes._subplots.AxesSubplot object at 0x1089026a0>,
                 <matplotlib.axes._subplots.AxesSubplot object at 0x108de6a90>]], dtype=object)
```

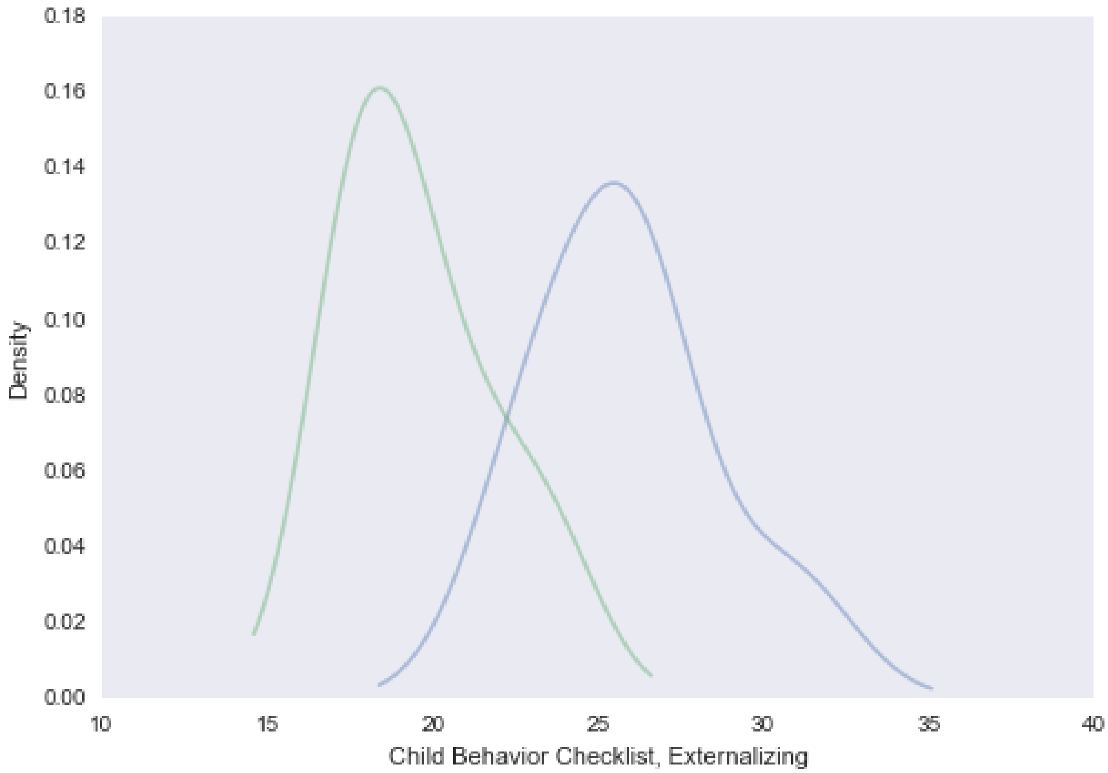


In [45]:
```
for x in analysis_subset.measure_instrument.unique():
    plt.figure()
    analysis_subset[analysis_subset.measure_instrument==x].baseline_mean.plot(kind='kde', alpha
    analysis_subset[analysis_subset.measure_instrument==x].end_treat_mean.plot(kind='kde', alph
    plt.gca().set_xlabel(x)
```




0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Density
−10
0
10
20
30
40
Eyberg Child Behaviour Inventory, Problem Subscale


0.12
0.10
0.08
0.06
0.04
0.02
0.00
Density
30
40
50
60
70
80
90
Child Behavior Checklist, Externalizing (T Score)



0.4 Meta-analysis

Number of studies in analysis subset

```
In [46]: unique_studies = analysis_subset.RefID.unique().tolist()
         len(unique_studies)
```

```
Out[46]: 26
```

We are restricting the analysis to the 4 most prevalent measure instruments in the database.

```
In [47]: unique_measures = analysis_subset.measure_instrument.unique().tolist()
         k = len(unique_measures)
         unique_measures
```

```
Out[47]: ['Eyberg Child Behaviour Inventory, Intensity Subscale',
          'Eyberg Child Behaviour Inventory, Problem Subscale',
          'Child Behavior Checklist, Externalizing (T Score)',
          'Child Behavior Checklist, Externalizing']
```

Three intervention components were coded:

- child_component
- parent_component

- multi_component

```
In [48]: analysis_subset[['child_only', 'parent_only', 'multi_component']].sum(0)
```

```
Out[48]: child_only          6
         parent_only        33
         multi_component    27
         dtype: int64
```

```
In [49]: p_male, age_cat, intvn = analysis_subset[['p_male', 'age_cat', 'intvn']].values.T
         child_only, parent_only, multi_component = analysis_subset[['child_only', 'parent_only',
                                                                      'multi_component']].valu

         change_n, change_mean, change_sd = analysis_subset[['eot_diff_n', 'eot_diff_mean', 'eot_diff_s
```

```
In [50]: school = (analysis_subset.age_cat.values==0).astype(int)
         pre_k = (analysis_subset.age_cat.values==1).astype(int)
         teen = (analysis_subset.age_cat.values==2).astype(int)
```

The response variable is a multivariate normal of dimension k=4, for each of the measure instruments.

$$\left(\begin{array}{c} m_1 \\ m_2 \\ m_3 \\ m_4 \end{array}\right)_i \sim \mathrm{MVN}(\mu, \Sigma)$$

Means for each study are a draw from a multivariate normal.

```
In [51]: wishart = True

         mu = pm.Normal('mu', 0, 0.001, value=[0]*k)

         if wishart:
             T = pm.Wishart('T', k, np.eye(k), value=np.eye(k))

             m = [pm.MvNormal('m_{}'.format(i), mu, T, value=[0]*k) for i in range(len(unique_studies))
         else:
             sigmas = pm.Uniform('sigmas', 0, 100, value=[10]*k)
             rhos = pm.Uniform('rhos', -1, 1, value=[0]*int(comb(k, 2)))

             Sigma = pm.Lambda('Sigma', lambda s=sigmas, r=rhos: np.array([[s[0]**2, s[0]*s[1]*r[0], s[
                                                                  [s[0]*s[1]*r[0], s[1]**2, s[1]*s[2]*r
                                                                  [s[0]*s[2]*r[1], s[1]*s[2]*r[3], s[2]
                                                                  [s[0]*s[3]*r[2], s[1]*s[3]*r[4], s[2]

             m = [pm.MvNormalCov('m_{}'.format(i), mu, Sigma, value=[0]*k) for i in range(len(unique_stu
```

Unique intervention labels for each component; we will use these for component random effects.

```
In [52]: unique_child_intvn = np.unique(intvn[child_only.astype(bool)]).tolist()
         unique_parent_intvn = np.unique(intvn[parent_only.astype(bool)]).tolist()
         unique_multi_intvn = np.unique(intvn[multi_component.astype(bool)]).tolist()
```

```
In [53]: # Indices to random effect labels
         child_component_index = [unique_child_intvn.index(x) for x in intvn[child_only.astype(bool)]]
         parent_component_index = [unique_parent_intvn.index(x) for x in intvn[parent_only.astype(bool)
         multi_component_index = [unique_multi_intvn.index(x) for x in intvn[multi_component.astype(boo
```

Treatment component random effects

$$X_i = \begin{bmatrix} x_c \\ x_p \\ x_f \end{bmatrix}_i$$

$$\beta_j^{(c)} \sim N(\mu_\beta^{(c)}, \tau_\beta^{(c)})$$

$$\beta_j^{(p)} \sim N(\mu_\beta^{(p)}, \tau_\beta^{(p)})$$

$$\beta_j^{(f)} \sim N(\mu_\beta^{(f)}, \tau_\beta^{(f)})$$

In [54]:
```
mu_beta = pm.Normal('mu_beta', 0, 0.001, value=[0]*3)
# sig_beta = pm.Uniform('sig_beta', 0, 100, value=1)
# tau_beta = sig_beta ** -2
tau_beta = pm.Gamma('tau_beta', 1, 0.1, value=1)

beta_c = pm.Normal('beta_c', mu_beta[0], tau_beta, value=[0]*len(unique_child_intvn))
beta_p = pm.Normal('beta_p', mu_beta[1], tau_beta, value=[0]*len(unique_parent_intvn))
beta_m = pm.Normal('beta_m', mu_beta[2], tau_beta, value=[0]*len(unique_multi_intvn))

b_c = pm.Lambda('b_c', lambda b=beta_c:
            np.array([b[unique_child_intvn.index(x)] if child_only[i] else 0 for i,x in enume
b_p = pm.Lambda('b_p', lambda b=beta_p:
            np.array([b[unique_parent_intvn.index(x)] if parent_only[i] else 0 for i,x in enum
b_m = pm.Lambda('b_m', lambda b=beta_m:
            np.array([b[unique_multi_intvn.index(x)] if multi_component[i] else 0 for i,x in e
```

In [55]:
```
best = pm.Lambda('best', lambda b=mu_beta:  (b==b.min()).astype(int))
```

Interaction of parent and multi-component with pre-k children.

In [56]:
```
interaction = False

if interaction:
    beta_pk_p = pm.Normal('beta_pk_p', 0, 1e-5, value=0)
    beta_pk_m = pm.Normal('beta_pk_m', 0, 1e-5, value=0)
    b_pk_p = pm.Lambda('b_pk_p', lambda b=beta_pk_p: b * parent_only * pre_k)
    b_pk_m = pm.Lambda('b_pk_m', lambda b=beta_pk_m: b * multi_component * pre_k)
```

In [57]:
```
betas = b_c + b_p + b_m

if interaction:
    betas = betas + b_pk_p + b_pk_m
```

Covariate effects of age and percent female.

$$\alpha \sim N(0, 1e5)$$

In [58]:
```
alpha_age = pm.Normal('alpha_age', 0, 1e-5, value=[1,2])
```

Unique study ID (`RefID`) and measure ID (`measure_instrument`) values.

In [59]:
```
study_id = [unique_studies.index(x) for x in analysis_subset.RefID]
measure_id = [unique_measures.index(x) for x in analysis_subset.measure_instrument]
```

Calculate expected response (treatment difference) as a function of treatment and covariates.

$$\theta_i = m_{j[i]k} + X_i\beta + \alpha x_{age}$$

In [60]:
```
baseline_sd = analysis_subset.baseline_sd.values

@pm.deterministic
def theta(m=m, betas=betas, alpha_age=alpha_age):

    mi = [m[i][j] for i,j in zip(study_id, measure_id)]

    age_effect = np.array([alpha_age[a-1] if a else 0 for a in age_cat])

    return(mi + baseline_sd*(betas + age_effect))
```

Expected treatment effect for pre-K undergoing multi-component intervention, measused by Eyberg Child Behaviour Inventory, Intensity Subscale

In [61]:
```
baseline = pm.MvNormalCov('baseline', mu, T, value=[0]*k)
```

In [63]:
```
ecbi_intensity_sd = baseline_sd[np.array(measure_id)==0].mean()

prek_intensity_pred = pm.Lambda('prek_intensity_pred',
                                lambda mu=baseline, a=alpha_age, b=mu_beta: mu[0] + ecbi_intensity
school_intensity_pred = pm.Lambda('school_intensity_pred',
                                lambda mu=baseline, a=alpha_age, b=mu_beta: mu[0] + ecbi_intensity
teen_intensity_pred = pm.Lambda('teen_intensity_pred',
                                lambda mu=baseline, a=alpha_age, b=mu_beta: mu[0] + ecbi_intensity

ecbi_problem_sd = baseline_sd[np.array(measure_id)==1].mean()

prek_problem_pred = pm.Lambda('prek_problem_pred',
                                lambda mu=baseline, a=alpha_age, b=mu_beta: mu[1] + ecbi_problem_s
school_problem_pred = pm.Lambda('school_problem_pred',
                                lambda mu=baseline, a=alpha_age, b=mu_beta: mu[1] + ecbi_problem_s
teen_problem_pred = pm.Lambda('teen_problem_pred',
                                lambda mu=baseline, a=alpha_age, b=mu_beta: mu[1] + ecbi_problem_s

cbct_sd = baseline_sd[np.array(measure_id)==2].mean()

prek_tscore_pred = pm.Lambda('prek_tscore_pred',
                                lambda mu=baseline, a=alpha_age, b=mu_beta: mu[2] + cbct_sd*(b + a
school_tscore_pred = pm.Lambda('school_tscore_pred',
                                lambda mu=baseline, b=mu_beta: mu[2] + cbct_sd*b )
teen_tscore_pred = pm.Lambda('teen_tscore_pred',
                                lambda mu=baseline, a=alpha_age, b=mu_beta: mu[2] + cbct_sd*(b + a

cbcr_sd = baseline_sd[np.array(measure_id)==3].mean()

prek_raw_pred = pm.Lambda('prek_raw_pred',
                                lambda mu=baseline, a=alpha_age, b=mu_beta: mu[3] + cbcr_sd*(b + a
school_raw_pred = pm.Lambda('school_raw_pred',
                                lambda mu=baseline, b=mu_beta: mu[3] + cbcr_sd*b )
teen_raw_pred = pm.Lambda('teen_raw_pred',
                                lambda mu=baseline, a=alpha_age, b=mu_beta: mu[3] + cbcr_sd*(b + a
```

Finally, the likelihood is just a normal distribution, with the observed standard error of the treatment effect as the standard deviation of the estimates.

$$d_i \sim N(\theta_i, \hat{\sigma}^2)$$

```
In [78]: change_se = change_sd/np.sqrt(change_n)
```

```
In [79]: d = pm.Normal('d', theta, change_se**-2, observed=True, value=change_mean)
```

Posterior predictive samples

```
In [80]: d_sim = pm.Normal('d_sim', theta, change_se**-2, size=len(change_mean))
```

```
In [81]: import appnope
         appnope.nope()

         M = pm.MCMC(locals())
         M.use_step_method(pm.AdaptiveMetropolis, [mu])
         M.use_step_method(pm.AdaptiveMetropolis, m)
         M.use_step_method(pm.AdaptiveMetropolis, mu_beta)
         M.use_step_method(pm.AdaptiveMetropolis, [beta_c, beta_p, beta_m])
```

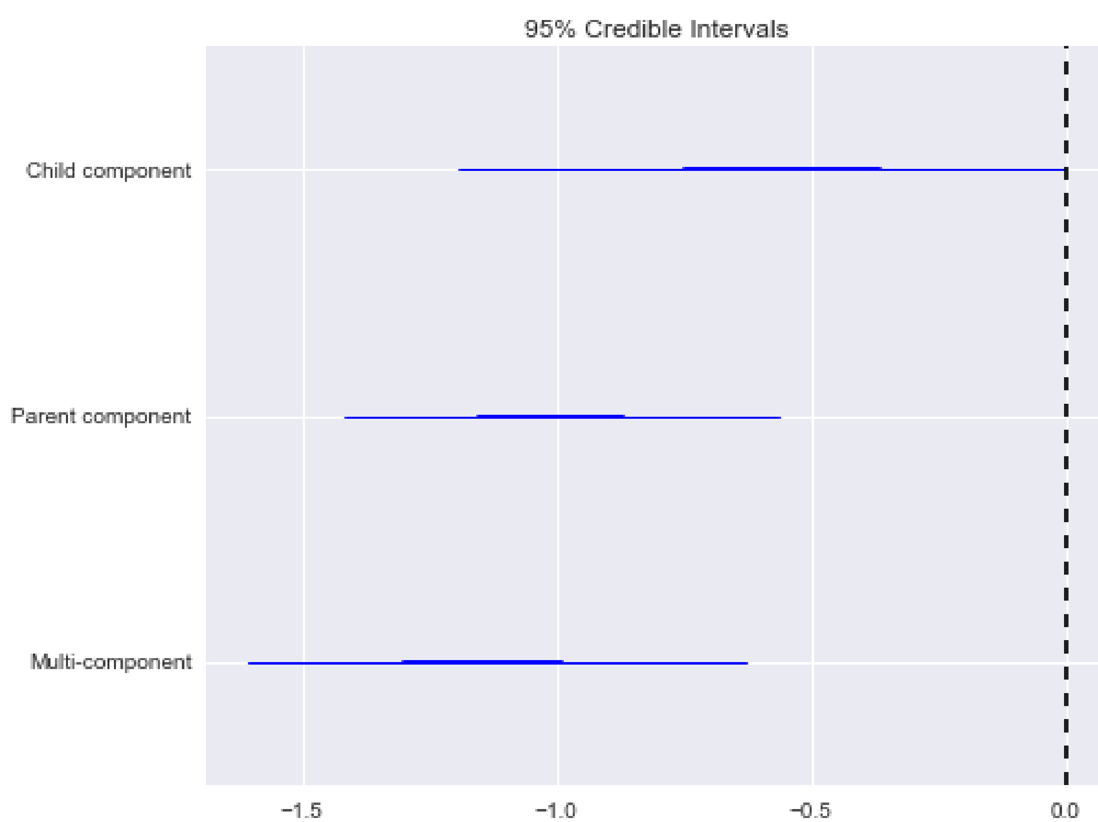
```
In [82]: M.sample(200000, 190000)
```

```
[-----------------100%-----------------] 200000 of 200000 complete in 585.7 sec
```

Summary of estimates of intervention components

```
In [83]: pm.Matplot.summary_plot([mu_beta], custom_labels=['Child component', 'Parent component',
                                                          'Multi-component'])
```

```
Could not calculate Gelman-Rubin statistics. Requires multiple chains of equal length.
```



```
In [84]: mu_beta.summary()
```

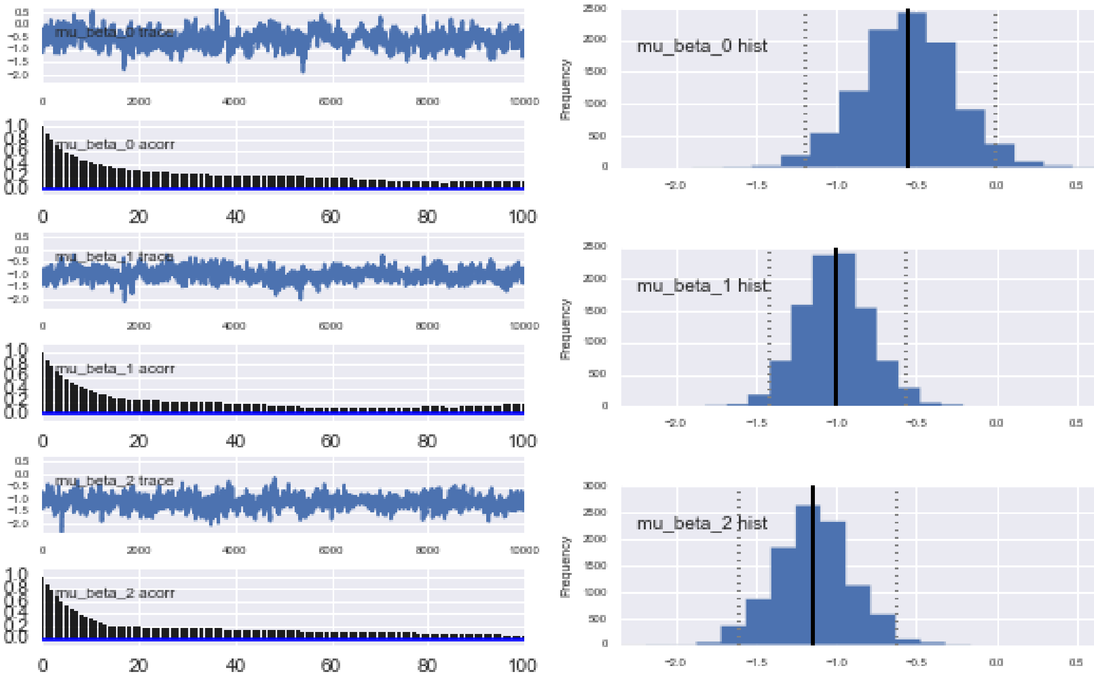
```
mu_beta:

        Mean             SD               MC Error         95% HPD interval
        -------------------------------------------------------------------
        -0.56            0.3              0.016              [-1.191 -0.004]
        -1.007           0.217            0.011              [-1.418 -0.564]
        -1.143           0.247            0.012              [-1.607 -0.629]


        Posterior quantiles:

        2.5             25              50              75             97.5
         |---------------|===============|===============|---------------|
        -1.165           -0.748          -0.558         -0.364        0.036
        -1.418           -1.155          -1.008         -0.868        -0.563
        -1.65            -1.302          -1.149         -0.989        -0.654
```

```
In [85]: best.summary()
```

```
best:

        Mean             SD               MC Error         95% HPD interval
        -------------------------------------------------------------------
        0.034            0.182            0.005                   [ 0.  0.]
        0.322            0.467            0.021                   [ 0.  1.]
        0.644            0.479            0.02                    [ 0.  1.]


        Posterior quantiles:

        2.5             25              50              75             97.5
         |---------------|===============|===============|---------------|
        0.0              0.0             0.0            0.0           1.0
        0.0              0.0             0.0            1.0           1.0
        0.0              0.0             1.0            1.0           1.0
```

```
In [86]: pm.Matplot.plot(mu_beta)
```

```
Plotting mu_beta_0
Plotting mu_beta_1
Plotting mu_beta_2
```

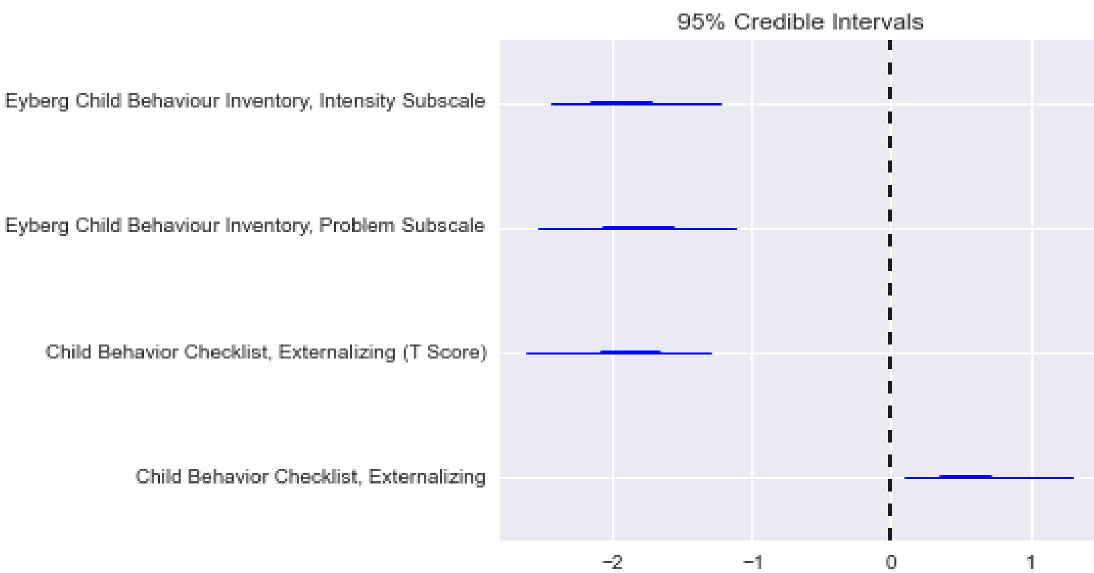



```
In [87]: if interaction:
             pm.Matplot.summary_plot([beta_pk_m, beta_pk_p])
```

Difference means by measure instrument.

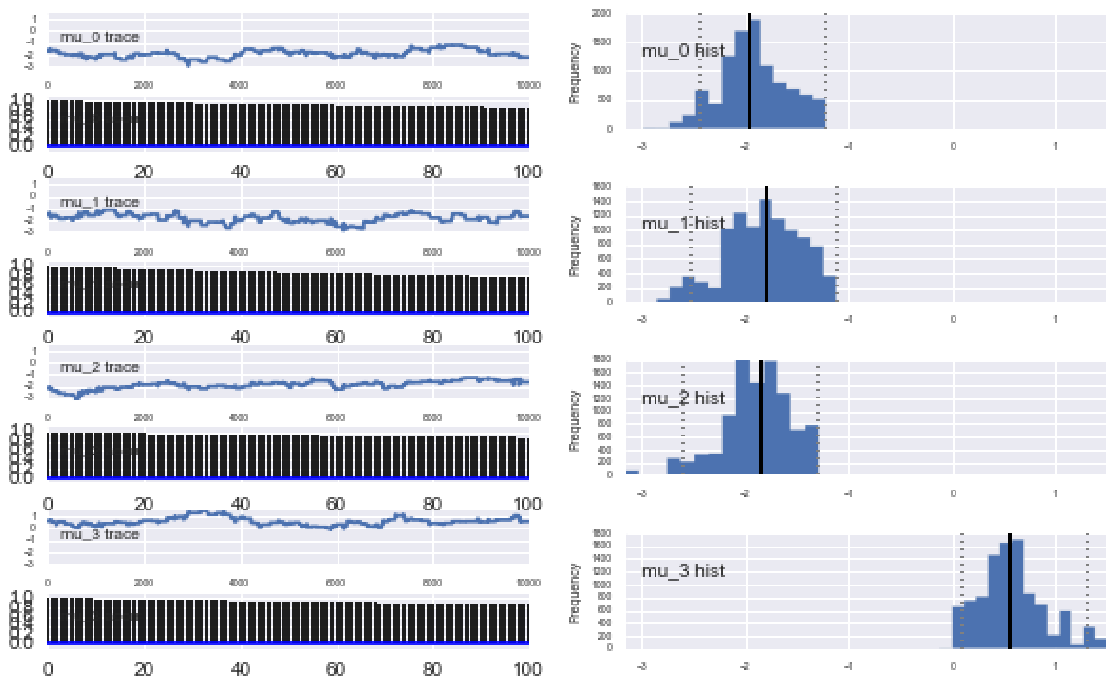
```
In [88]: plt.figure(figsize=(24,4))
         pm.Matplot.summary_plot([mu], custom_labels=unique_measures)
```

```
Could not calculate Gelman-Rubin statistics. Requires multiple chains of equal length.
```



```
In [89]: pm.Matplot.plot(mu)
```

```
Plotting mu_0
Plotting mu_1
Plotting mu_2
Plotting mu_3
```

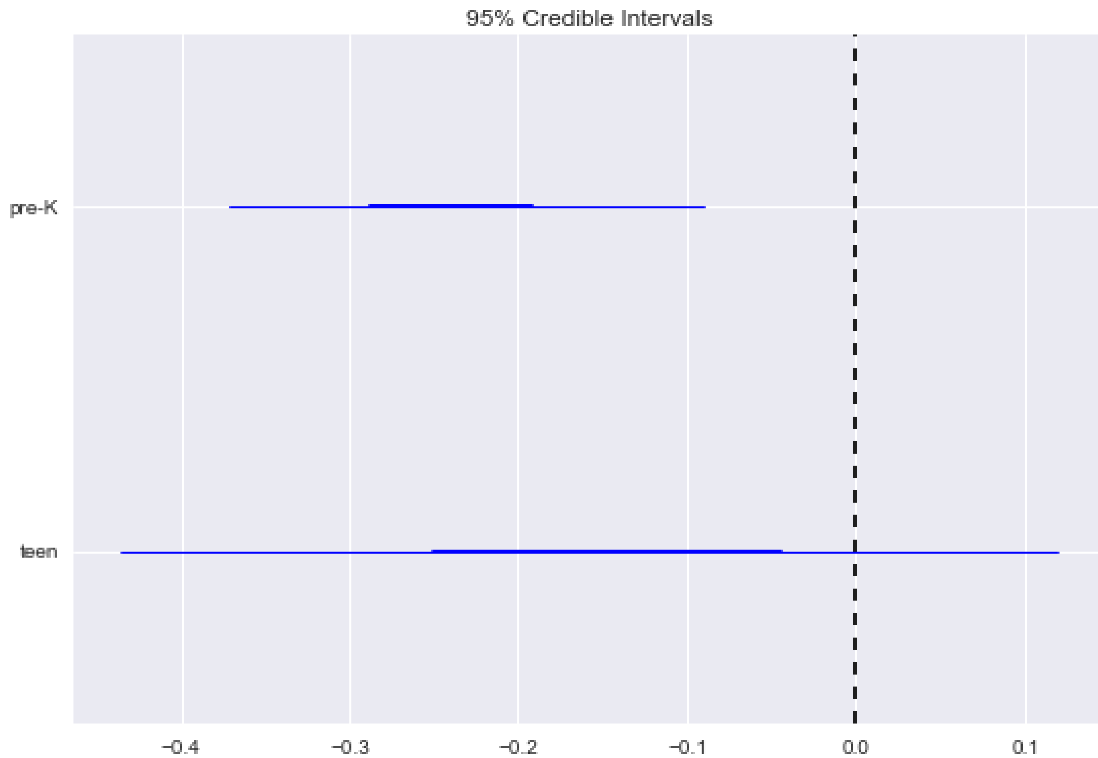



```
In [90]: if not wishart:
             plt.figure(figsize=(24,4))
             pm.Matplot.summary_plot([sigmas], custom_labels=unique_measures)

In [91]: if not wishart:
             pm.Matplot.summary_plot([rhos], custom_labels=['rho12', 'rho13', 'rho14', 'rho23', 'rho24'
```

Age effects for pre-k (top) and teen (bottom) groups, relative to pre-teen.

```
In [92]: pm.Matplot.summary_plot([alpha_age], custom_labels=['pre-K', 'teen'])
```

```
Could not calculate Gelman-Rubin statistics. Requires multiple chains of equal length.
```

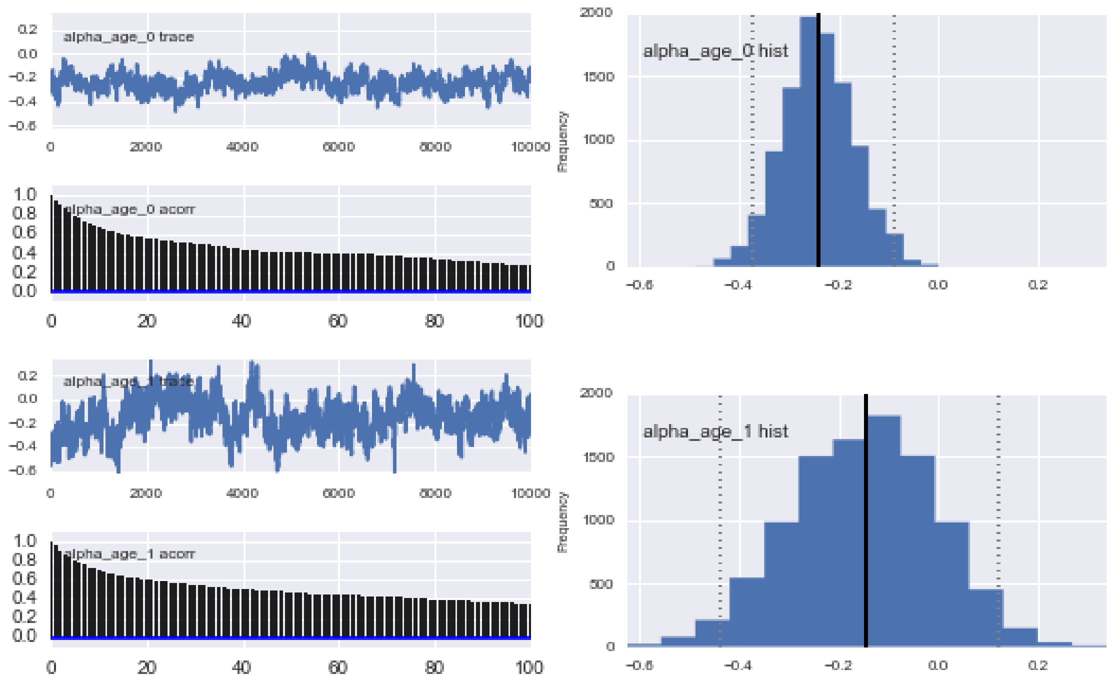


In [93]: alpha_age.summary()

```
alpha_age:

        Mean             SD               MC Error        95% HPD interval
        ------------------------------------------------------------------
        -0.241           0.072            0.005            [-0.372 -0.09 ]
        -0.149           0.145            0.011            [-0.436  0.12 ]


        Posterior quantiles:

        2.5             25              50              75             97.5
         |---------------|===============|===============|---------------|
        -0.379           -0.288          -0.242          -0.192         -0.097
        -0.436           -0.251          -0.145          -0.044         0.12
```

In [94]: pm.Matplot.plot(alpha_age)

```
Plotting alpha_age_0
Plotting alpha_age_1
```



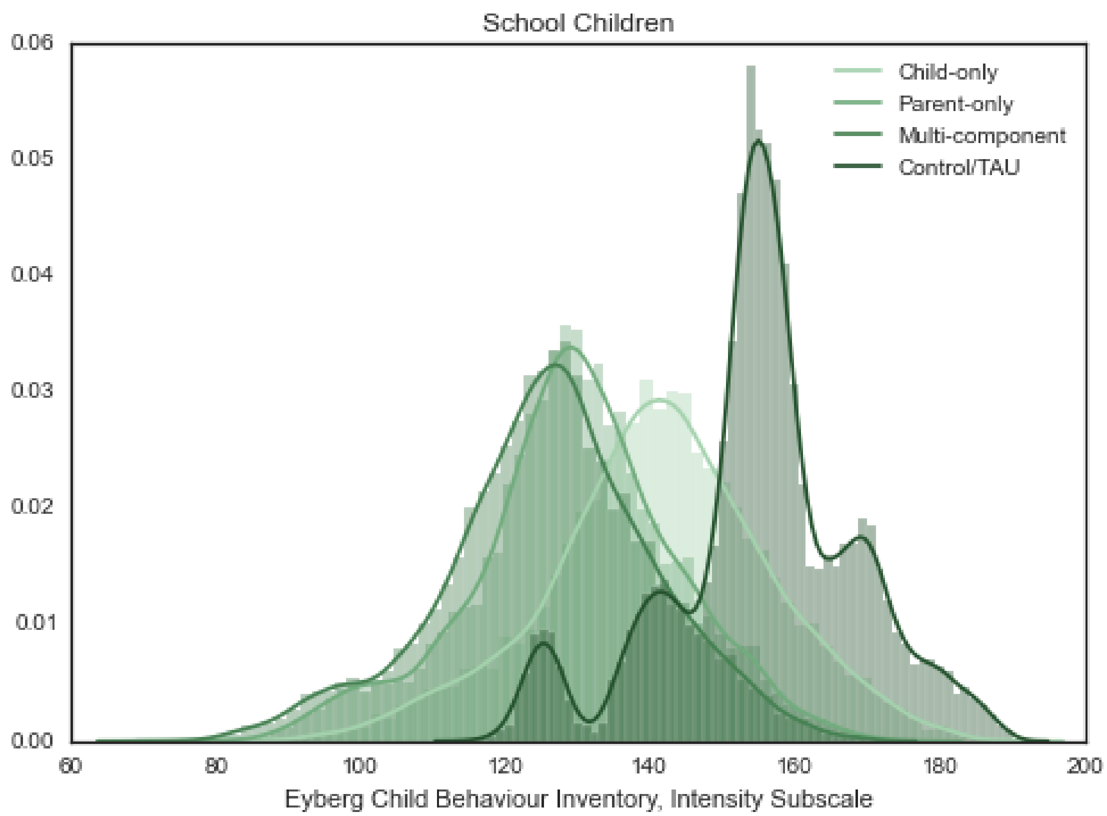
0.5 Outcome Plots

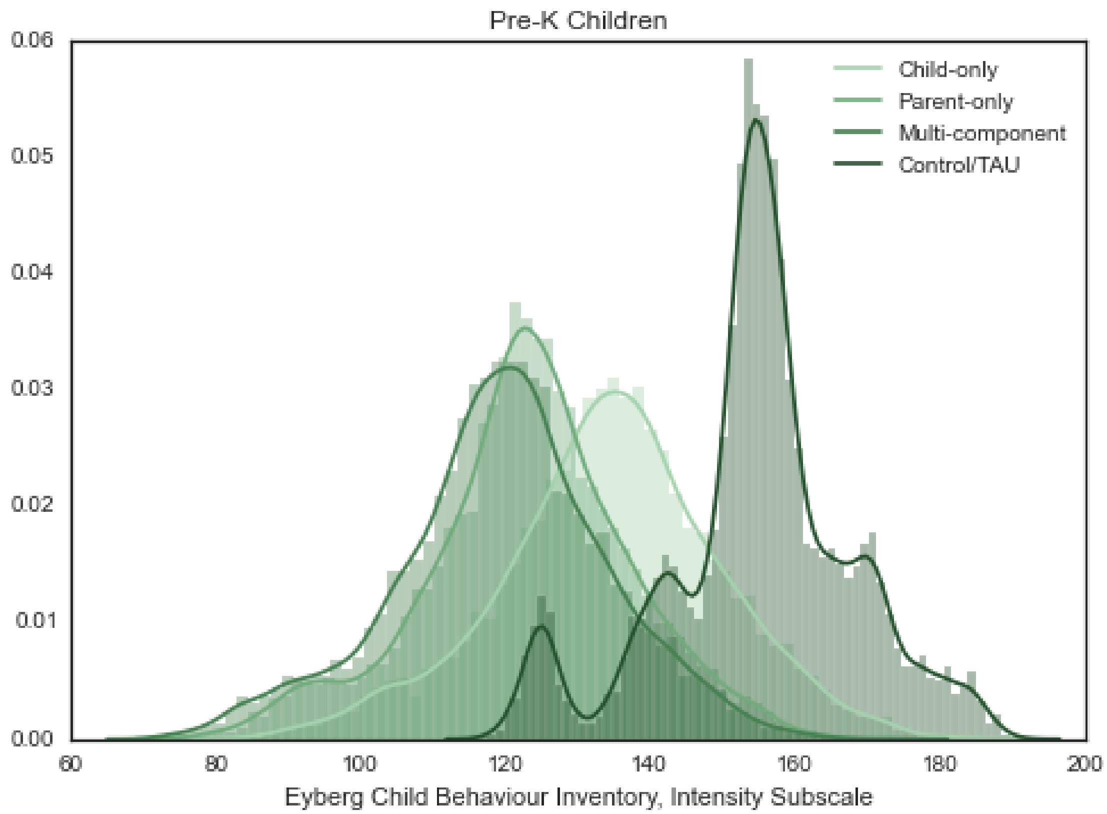
In [95]:
```
traces = [[school_intensity_pred, prek_intensity_pred, teen_intensity_pred],
          [school_problem_pred, prek_problem_pred, teen_problem_pred],
          [school_tscore_pred, prek_tscore_pred, teen_tscore_pred],
          [school_raw_pred, prek_raw_pred, teen_raw_pred]]
```

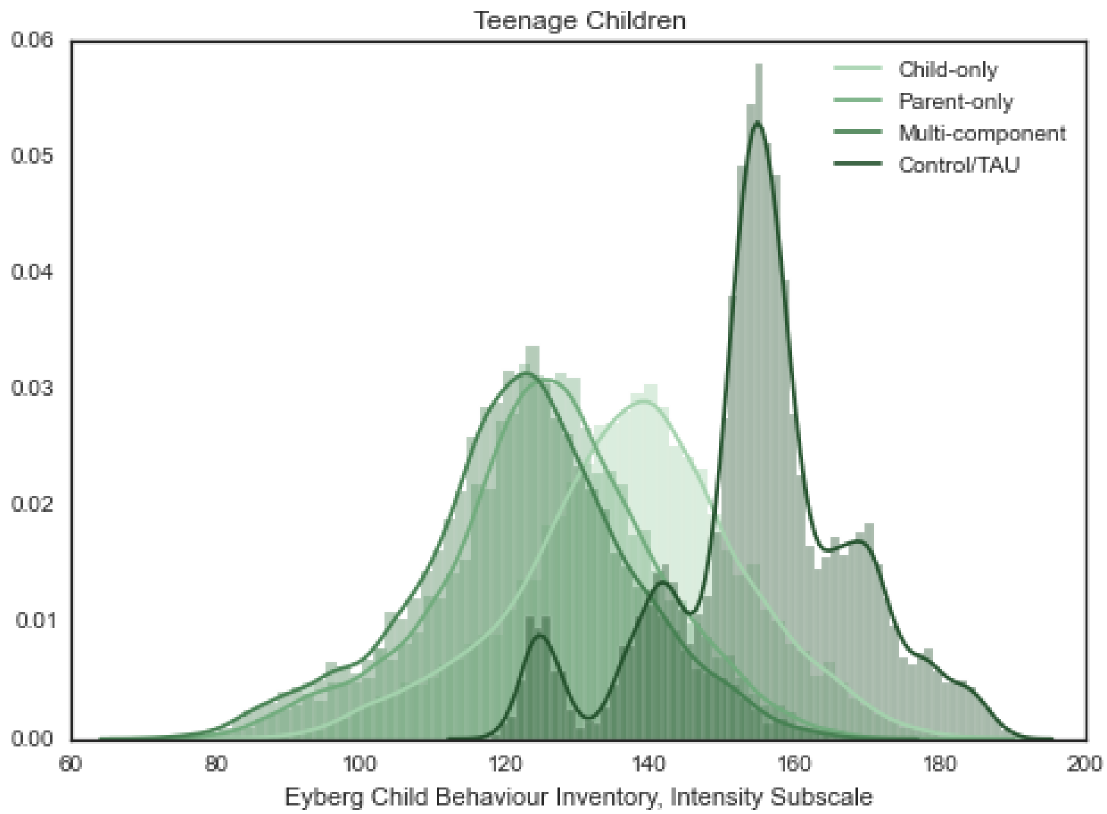
In [96]:
```
sb.set(style="white", palette="hot")
sb.despine(left=True)

#colors = '#fef0d9','#fdcc8a','#fc8d59','#d7301f'
colors = sb.cubehelix_palette(4, start=2, rot=0, dark=.25, light=.75, reverse=False)

titles = ['School Children', 'Pre-K Children', 'Teenage Children']

for i,measure in enumerate(unique_measures):

    measure_traces = traces[i]

    for j, trace in enumerate(measure_traces):

        x = np.random.choice(analysis_subset[analysis_subset.measure_instrument==measure].basel

        c1, p1, m1 = trace.trace().T

        plt.figure()
        g = sb.distplot(x + c1, color=colors[0])
        sb.distplot(x + p1, color=colors[1])
        sb.distplot(x + m1, color=colors[2])
```

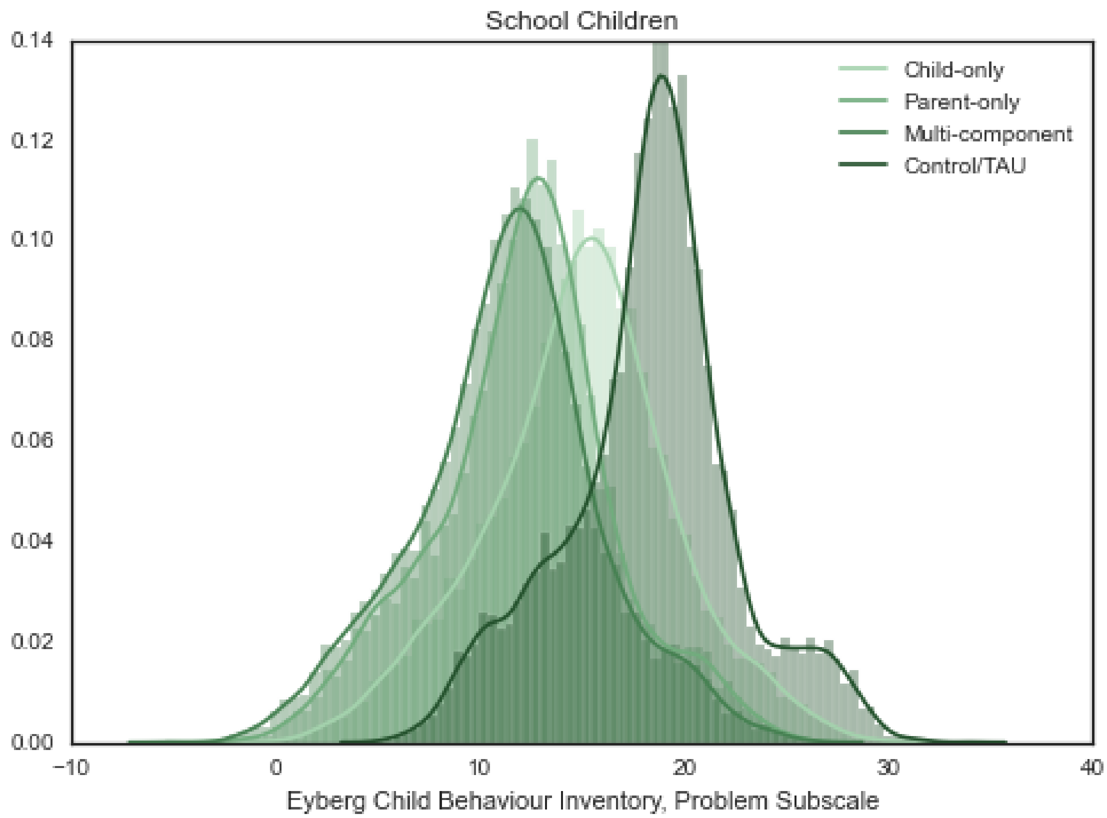
```
            if j:
                age_effect = alpha_age.trace()[:, j-1]
            else:
                age_effect = 0
            sb.distplot(x + baseline.trace()[:, i] + age_effect, color=colors[3]);
            g.set_title(titles[j])

            g.legend(g.lines, ['Child-only', 'Parent-only', 'Multi-component', 'Control/TAU'])

            g.set_xlabel(measure)
```

```
<matplotlib.figure.Figure at 0x108a7e4a8>
```

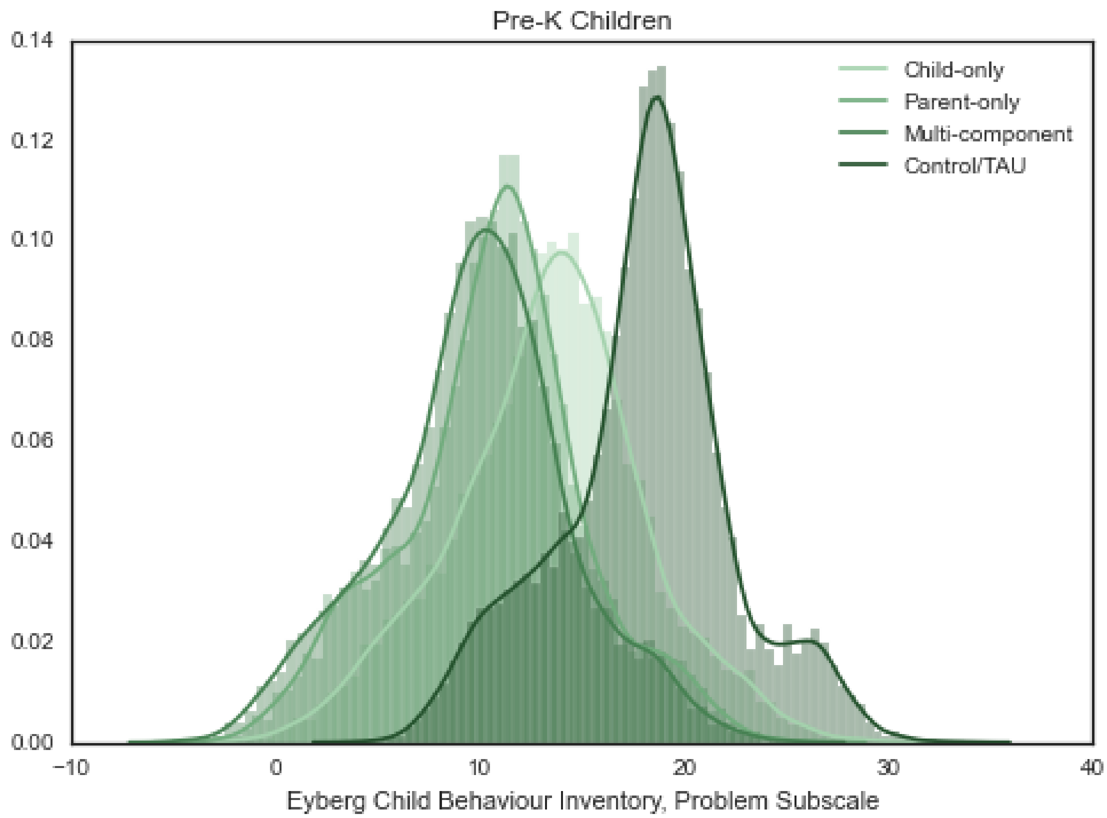


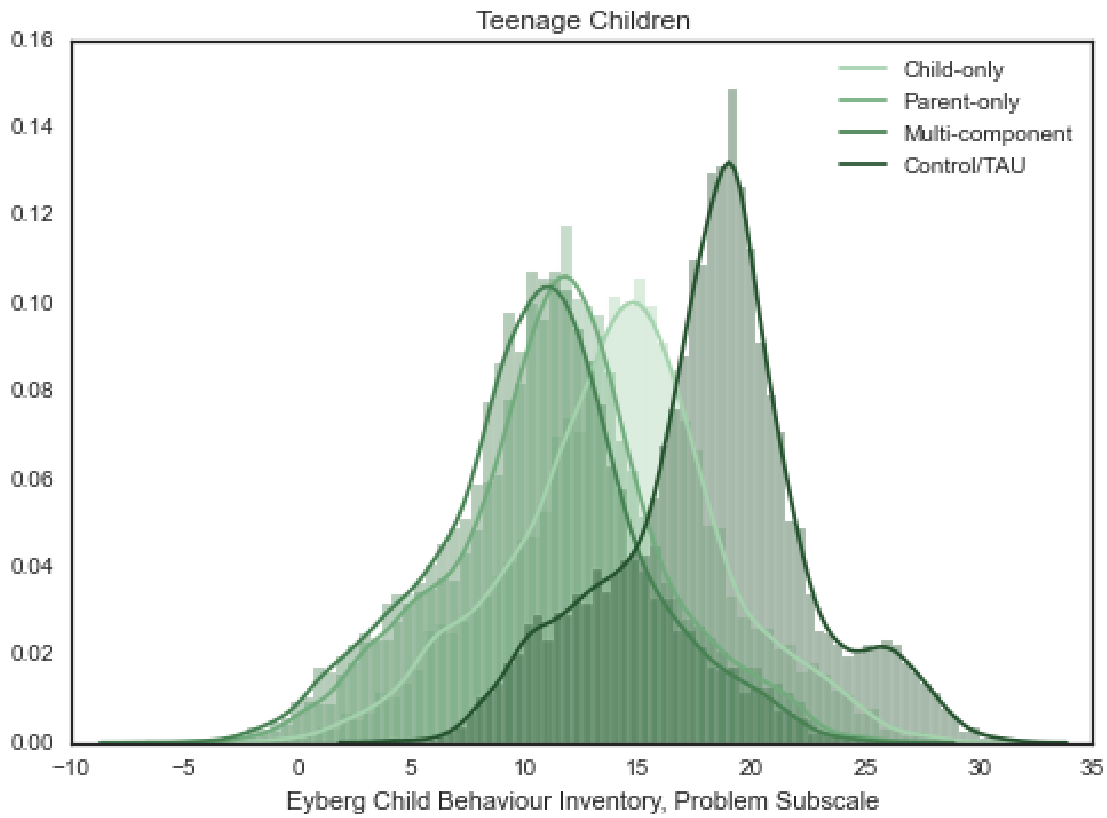

Pre-K Children
Child-only
Parent-only
Multi-component
Control/TAU
0.06
0.05
0.04
0.03
0.02
0.01
0.00
60
80
100
120
140
160
180
200
Eyberg Child Behaviour Inventory, Intensity Subscale

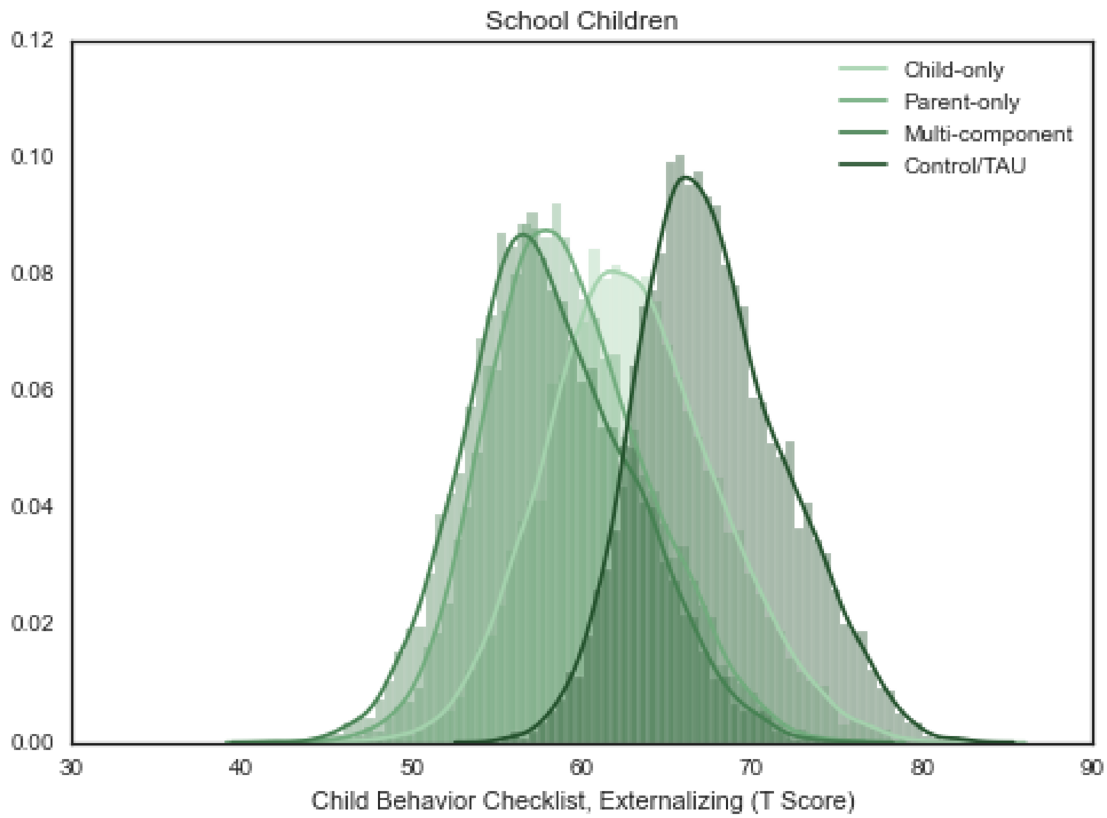

Teenage Children
0.06
0.05
0.04
0.03
0.02
0.01
0.00
60
80
100
120
140
160
180
200
Eyberg Child Behaviour Inventory, Intensity Subscale
Child-only
Parent-only
Multi-component
Control/TAU

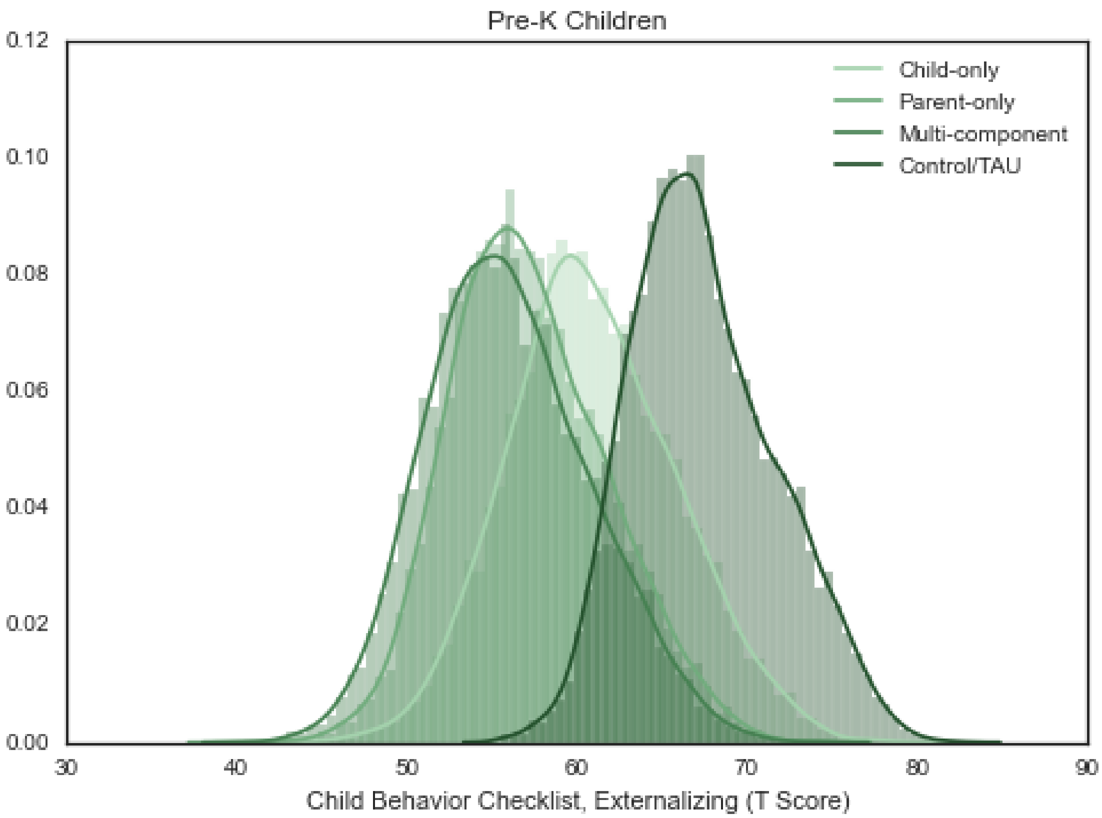


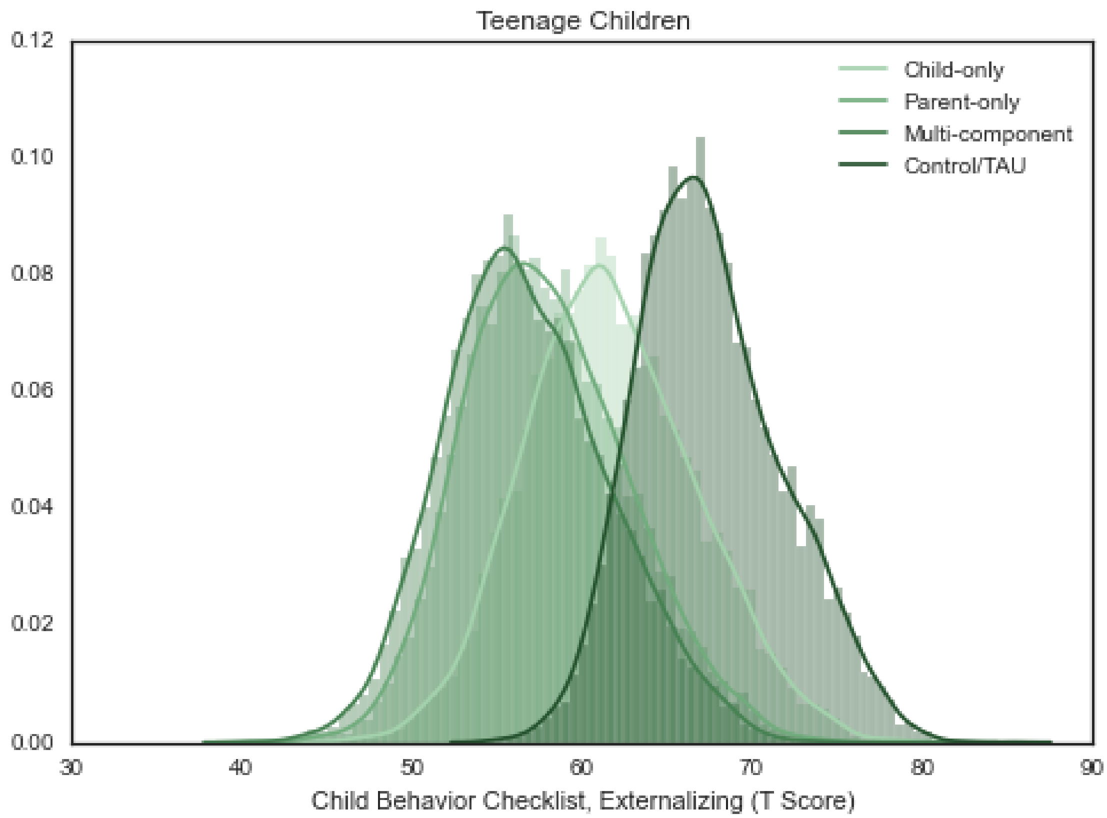
School Children
Child-only
Parent-only
Multi-component
Control/TAU
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
−10
0
10
20
30
40
Eyberg Child Behaviour Inventory, Problem Subscale

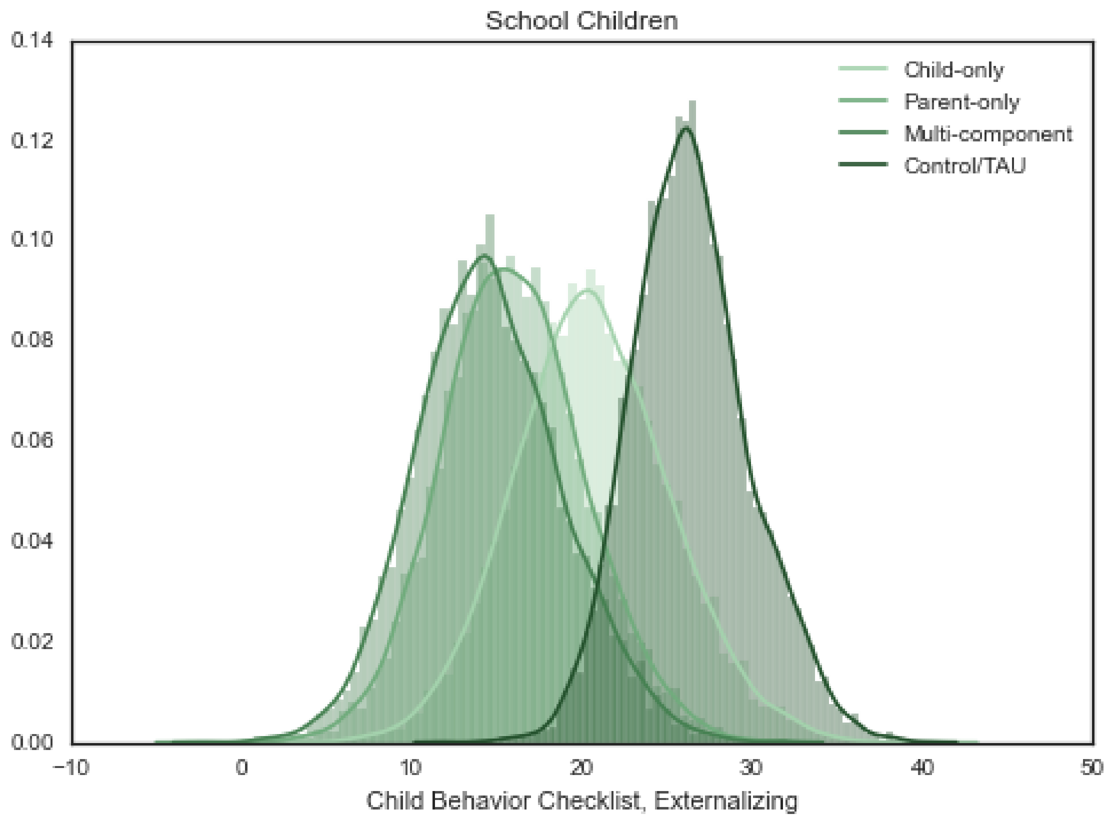

Pre-K Children
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
−10
0
10
20
30
40
Child-only
Parent-only
Multi-component
Control/TAU
Eyberg Child Behaviour Inventory, Problem Subscale

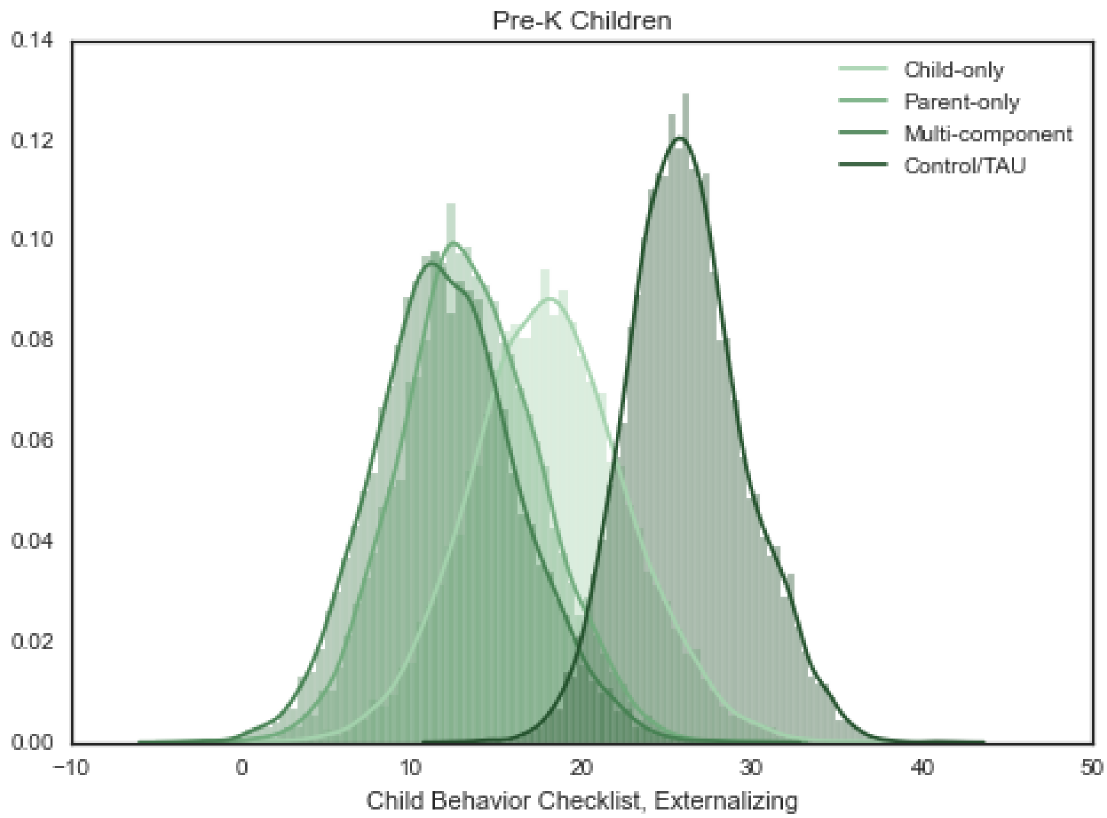

Teenage Children
Child-only
Parent-only
Multi-component
Control/TAU
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
0.16
−10
−5
0
5
10
15
20
25
30
35
Eyberg Child Behaviour Inventory, Problem Subscale

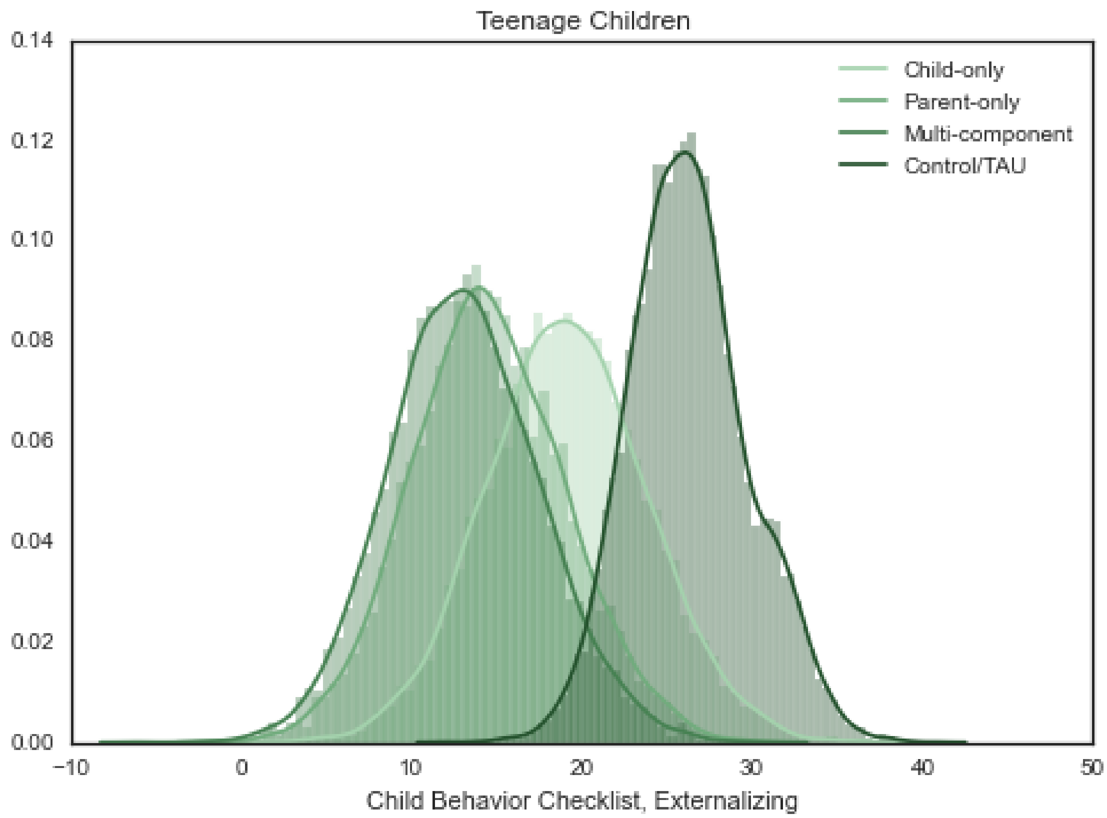

School Children
0.12
0.10
0.08
0.06
0.04
0.02
0.00
30
40
50
60
70
80
90
Child Behavior Checklist, Externalizing (T Score)
Child-only
Parent-only
Multi-component
Control/TAU


Pre-K Children
Child-only
Parent-only
Multi-component
Control/TAU
0.12
0.10
0.08
0.06
0.04
0.02
0.00
30
40
50
60
70
80
90
Child Behavior Checklist, Externalizing (T Score)


Teenage Children
Child-only
Parent-only
Multi-component
Control/TAU
0.12
0.10
0.08
0.06
0.04
0.02
0.00
30
40
50
60
70
80
90
Child Behavior Checklist, Externalizing (T Score)


School Children
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
−10
0
10
20
30
40
50
Child Behavior Checklist, Externalizing
Child-only
Parent-only
Multi-component
Control/TAU


Pre-K Children
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
−10
0
10
20
30
40
50
Child Behavior Checklist, Externalizing
Child-only
Parent-only
Multi-component
Control/TAU




```
In [97]: x1 = np.random.choice(analysis_subset[analysis_subset.measure_instrument==unique_measures[0]].l
```

```
In []: #pm.Matplot.gof_plot(d_sim, change_mean, verbose=0)
```