In [3]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [4]:

```
# Distribution within Categories
```

In [5]:

```
"""
Boxplot
Violinplot
Swarmplot
Boxenplot
"""
```

Out[5]:

```
'\nBoxplot\nViolinplot\nSwarmplot\nBoxenplot\n'
```

In [9]:

```
df= pd.read_csv("StudentsPerformance.csv")
df
```

Out[9]:

| | gender | race/ethnicity | parental level of education | lunch | test preparation course | math score | reading score | writing score |
|---|---|---|---|---|---|---|---|---|
| **0** | female | group B | bachelor's degree | standard | none | 72 | 72 | 74 |
| **1** | female | group C | some college | standard | completed | 69 | 90 | 88 |
| **2** | female | group B | master's degree | standard | none | 90 | 95 | 93 |
| **3** | male | group A | associate's degree | free/reduced | none | 47 | 57 | 44 |
| **4** | male | group C | some college | standard | none | 76 | 78 | 75 |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... |
| **995** | female | group E | master's degree | standard | completed | 88 | 99 | 95 |
| **996** | male | group C | high school | free/reduced | none | 62 | 55 | 55 |
| **997** | female | group C | high school | free/reduced | completed | 59 | 71 | 65 |
| **998** | female | group D | some college | standard | completed | 68 | 78 | 77 |
| **999** | female | group D | some college | free/reduced | none | 77 | 86 | 86 |

1000 rows × 8 columns

In [10]:

```
df.head()
```

Out[10]:

| | gender | race/ethnicity | parental level of education | lunch | test preparation course | math score | reading score | writing score |
|---|---|---|---|---|---|---|---|---|
| **0** | female | group B | bachelor's degree | standard | none | 72 | 72 | 74 |
| **1** | female | group C | some college | standard | completed | 69 | 90 | 88 |
| **2** | female | group B | master's degree | standard | none | 90 | 95 | 93 |
| **3** | male | group A | associate's degree | free/reduced | none | 47 | 57 | 44 |
| **4** | male | group C | some college | standard | none | 76 | 78 | 75 |

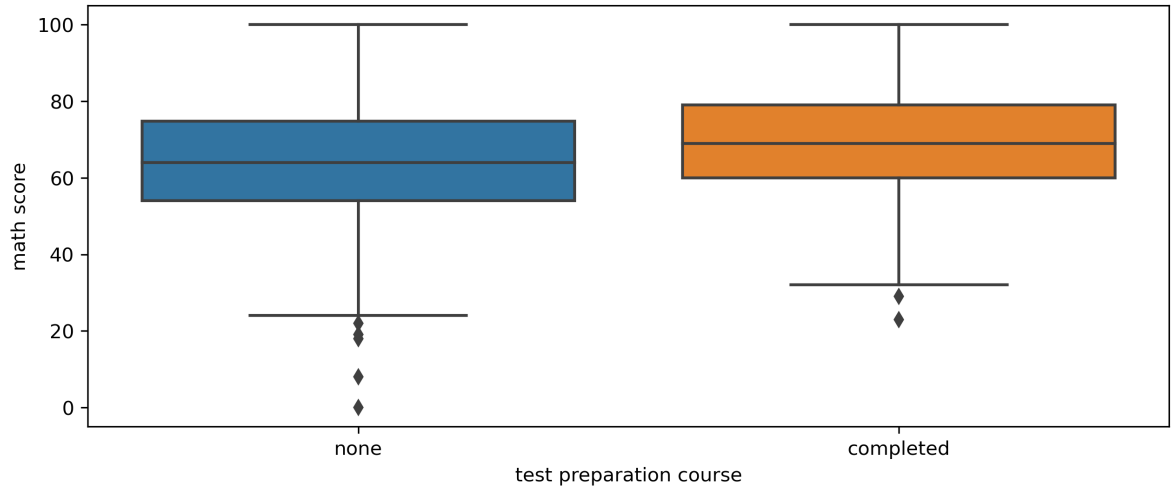
# boxplot

In [12]:

```
plt.figure(figsize=(10,4),dpi=300)
sns.boxplot(data=df,y="math score",x="test preparation course")
```

Out[12]:

```
<AxesSubplot:xlabel='test preparation course', ylabel='math score'>
```

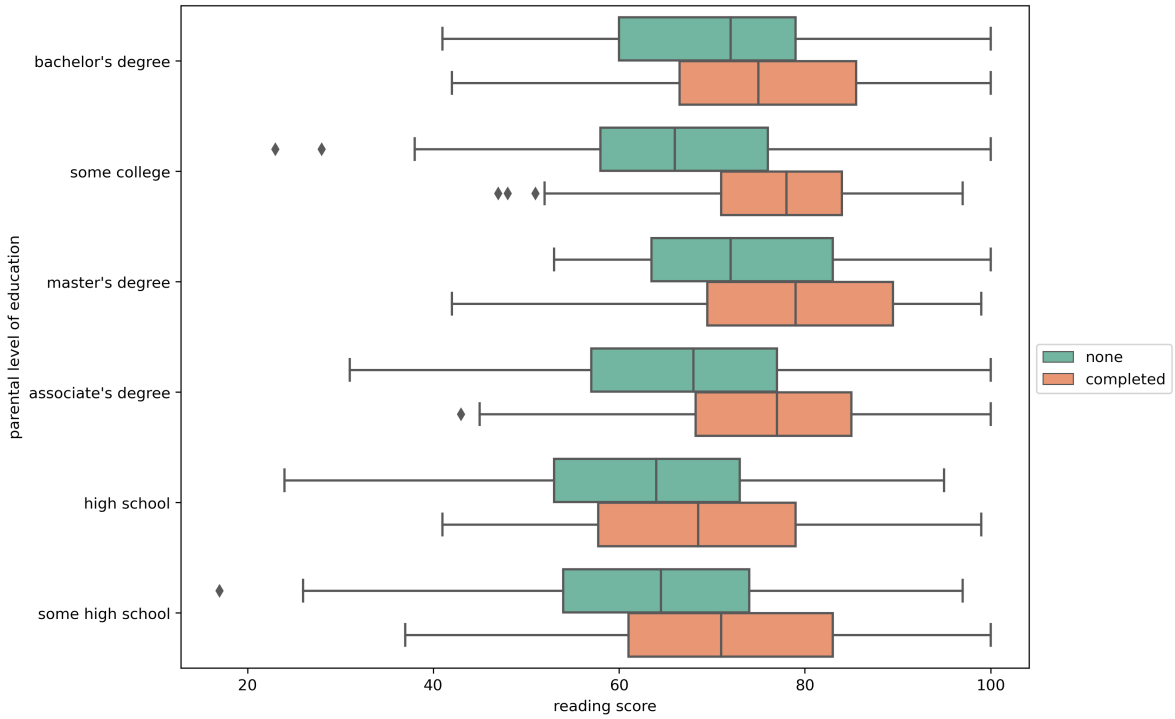


In [27]:

```
plt.figure(figsize=(10,8),dpi=300)
sns.boxplot(data=df,x="reading score",y="parental level of education",
            hue="test preparation course",palette="Set2")
plt.legend(bbox_to_anchor=(1,0.5))
```

Out[27]:

```
<matplotlib.legend.Legend at 0x1d4d20bbcd0>
```



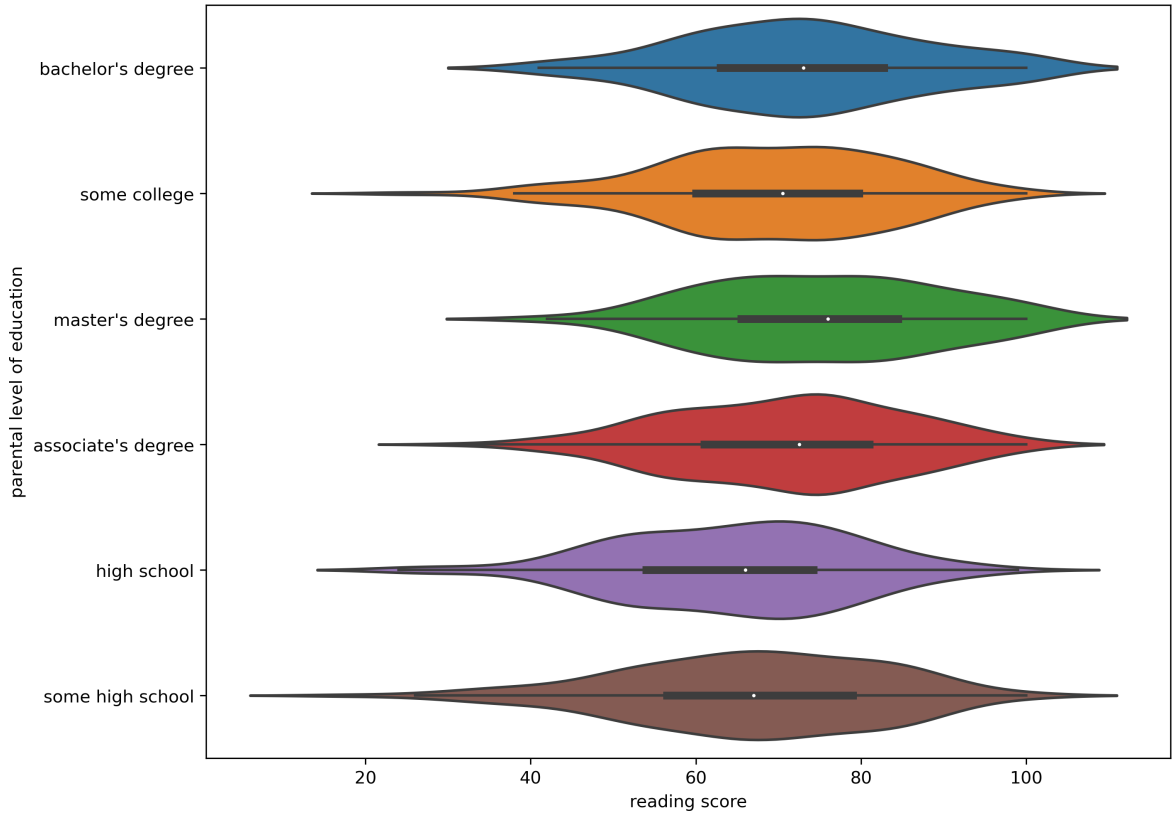

# violinplot

In [32]:

```
plt.figure(figsize=(10,8),dpi=300)
sns.violinplot(data=df,x="reading score",y="parental level of education"
               )
#plt.legend(bbox_to_anchor=(1,0.5))
```

Out[32]:

```
<AxesSubplot:xlabel='reading score', ylabel='parental level of education'>
```

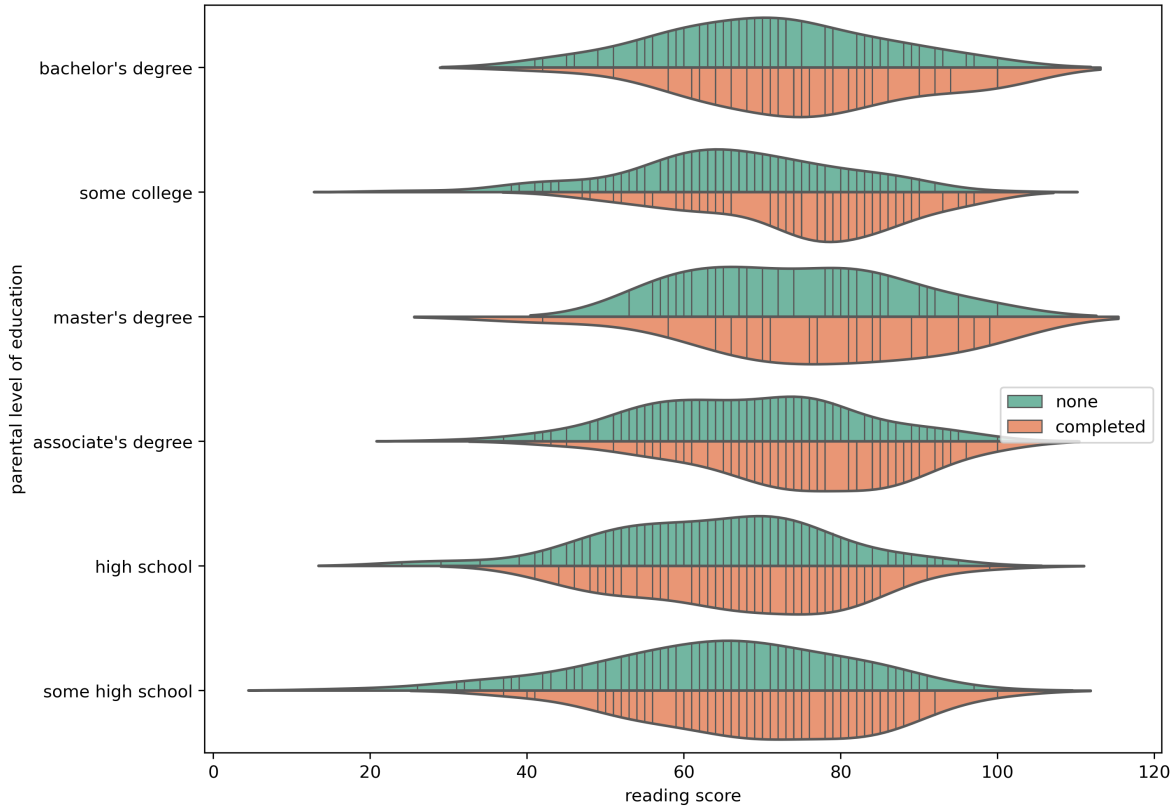


In [36]:

```
plt.figure(figsize=(10,8),dpi=300)
sns.violinplot(data=df,x="reading score",y="parental level of education",
           hue="test preparation course",palette="Set2",
              split=True,inner="stick")#inner="quartile"
plt.legend(bbox_to_anchor=(1,0.5))
```

Out[36]:

```
<matplotlib.legend.Legend at 0x1d4d34d32b0>
```

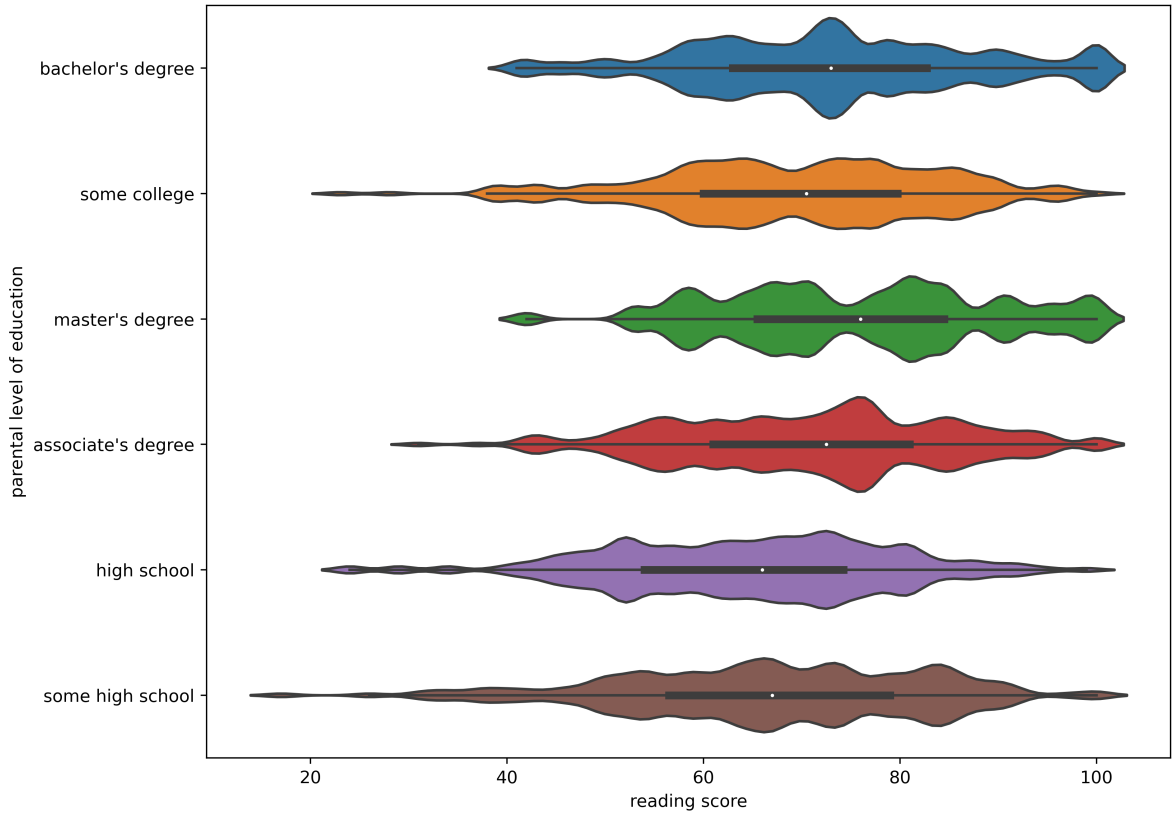


In [39]:

```
plt.figure(figsize=(10,8),dpi=400)
sns.violinplot(data=df,x="reading score",y="parental level of education",
         bw=0.1)
#plt.legend(bbox_to_anchor=(1,0.5))
```

Out[39]:

```
<AxesSubplot:xlabel='reading score', ylabel='parental level of education'>
```



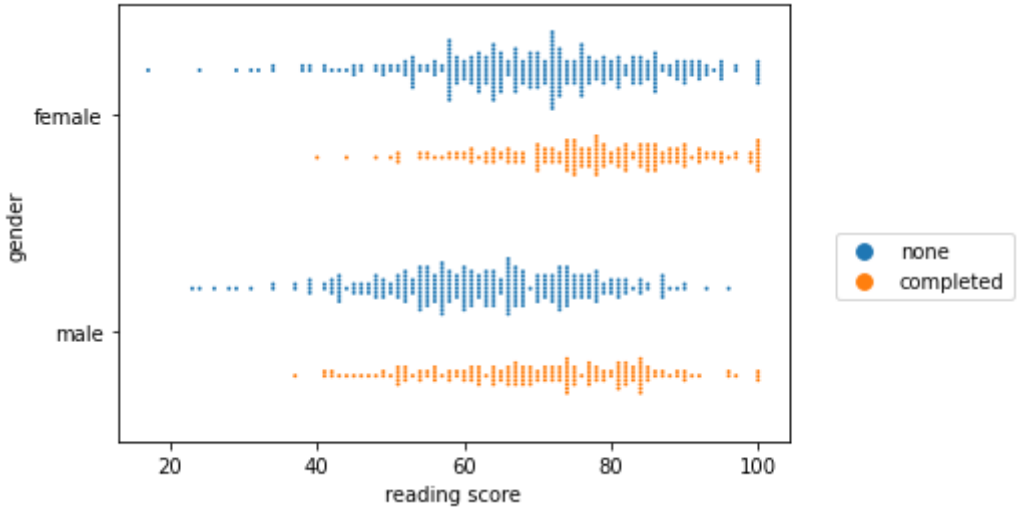

# swarmplot

In [48]:

```
sns.swarmplot(data=df,x="reading score",size=2,y="gender",
              hue="test preparation course",dodge=True)
plt.legend(bbox_to_anchor=(1.35,0.5))
```

Out[48]:

```
<matplotlib.legend.Legend at 0x1d4ccf85640>
```



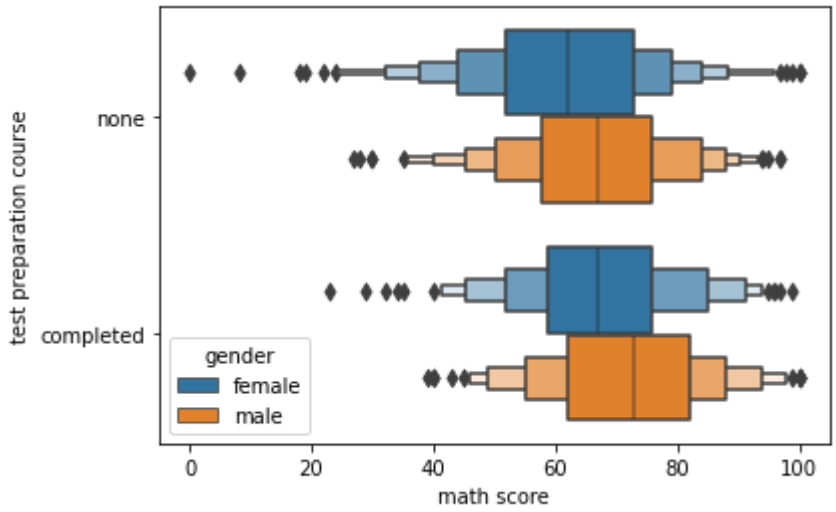

# boxenplot

In [51]:

```
sns.boxenplot(x="math score",y="test preparation course",data=df,hue="gender")
```

Out[51]:

```
<AxesSubplot:xlabel='math score', ylabel='test preparation course'>
```




In [ ]:

```

```