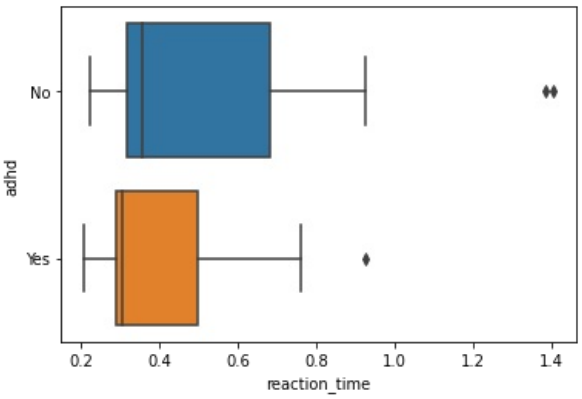
In [9]:
```
import pandas as pd
import scipy.stats as stats
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
```

In [10]:
```
DATA_SOURCE = "./data/reaction-time-data-5.csv"
print(pd.read_csv(DATA_SOURCE).head())
```

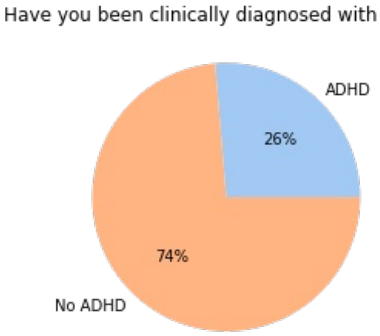
```
                 time    age       grade adhd currently_medicated  \
0  4/26/2022 14:10:40  14-16  Sophomore   No                 NaN
1  4/26/2022 14:11:10  17-18      Senior  Yes                  No
2  4/26/2022 14:13:46  17-18      Junior   No                 NaN
3  4/26/2022 14:20:28  17-18      Senior   No                 NaN
4  4/26/2022 14:20:29  17-18      Senior   No                 NaN

   reaction_time
0         0.9222
1         0.3496
2         0.7480
3         0.6458
4         0.6702
```

In [24]:
```
    """Stacked box plot of reaction time, both with and without ADHD."""
    df = pd.read_csv(DATA_SOURCE)

    ADHD = df.loc[df["adhd"] == "Yes"]
    NON_ADHD = df.loc[df["adhd"] == "No"]

    concatenated = pd.concat(
        [NON_ADHD.assign(dataset="No ADHD"), ADHD.assign(dataset="ADHD")]
    )

    sns.boxplot(y="adhd", x="reaction_time", data=concatenated)
    plt.savefig('view_data_adhd.png', dpi=1000)
    plt.show()
```



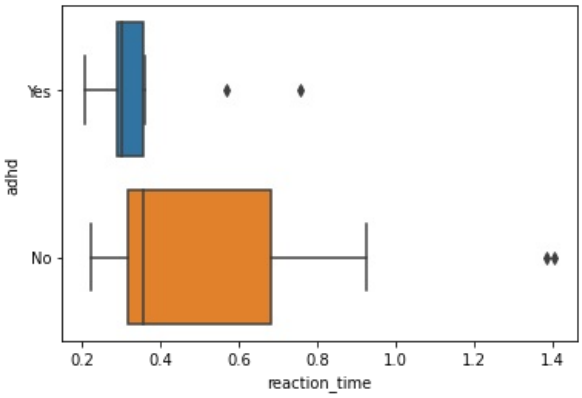
In [25]:
```
    """Pie chart of people with and without ADHD."""
    df = pd.read_csv(DATA_SOURCE)

    ADHD = len(df.loc[df["adhd"] == "Yes"])
    NON_ADHD = len(df.loc[df["adhd"] == "No"])

    colors = sns.color_palette("pastel")[0:5]

    plt.pie(
        [ADHD, NON_ADHD], labels=["ADHD", "No ADHD"], colors=colors, autopct="%.0f%%"
    )
    plt.title("Have you been clinically diagnosed with ADHD?")
    plt.savefig('view_data_adhd_pie.png', dpi=1000)
```

```
plt.show()
```



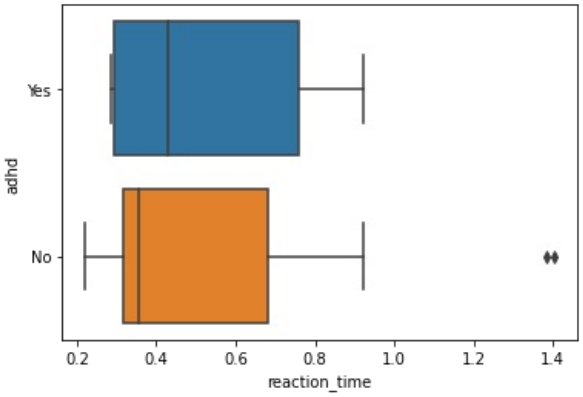
In [26]:

```
"""Box plot of reaction time,
of (people with adhd who were not on medication) and (people who do not have adhd).
"""
df = pd.read_csv(DATA_SOURCE)

NON_MED = df.loc[df["currently_medicated"] == "No"]
NON_ADHD = df.loc[df["adhd"] == "No"]

concatenated = pd.concat(
    [NON_MED.assign(dataset="No MED"), NON_ADHD.assign(dataset="No ADHD")]
)
sns.boxplot(y="adhd", x="reaction_time", data=concatenated)

plt.savefig('view_data_no_med_vs_no_adhd.png', dpi=1000)
plt.show()
```



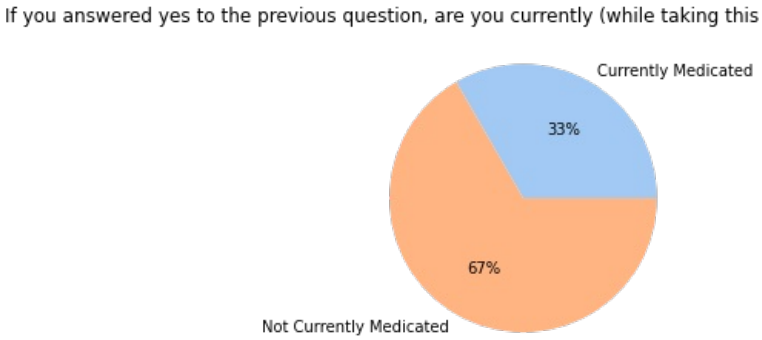
In [27]:

```
"""Box plot of reaction time,
of (people with adhd who were on medication) and (people who do not have adhd).
"""
df = pd.read_csv(DATA_SOURCE)

NON_MED = df.loc[df["currently_medicated"] == "Yes"]
NON_ADHD = df.loc[df["adhd"] == "No"]

concatenated = pd.concat(
    [NON_MED.assign(dataset="No MED"), NON_ADHD.assign(dataset="No ADHD")]
)
sns.boxplot(y="adhd", x="reaction_time", data=concatenated)

plt.savefig('view_data_yes_med_vs_no_adhd.png', dpi=1000)
plt.show()
```



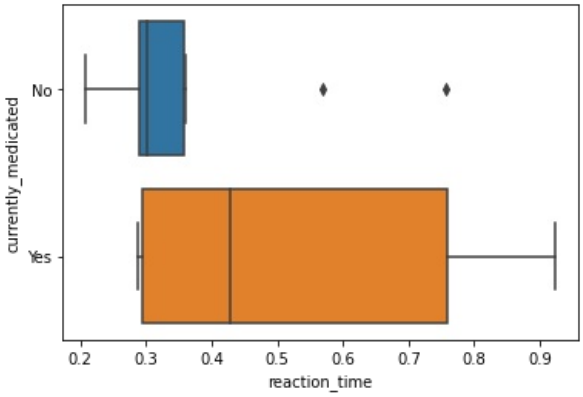
In [28]:

```
    """Pie chart of people with ADHD, who are medicated vs who are not."""
    df = pd.read_csv(DATA_SOURCE)

    MED = len(df.loc[df["currently_medicated"] == "Yes"])
    NON_MED = len(df.loc[df["currently_medicated"] == "No"])

    colors = sns.color_palette("pastel")[0:5]

    plt.pie(
        [MED, NON_MED],
        labels=["Currently Medicated", "Not Currently Medicated"],
        colors=colors,
        autopct="%.0f%%",
    )
    plt.title(
        "If you answered yes to the previous question, are you currently (while taking this
survey) medicated for ADHD?"
    )
    plt.savefig('view_data_med_pie.png', dpi=1000)
    plt.show()
```



In [29]:

```
    """Box plot of reaction time of people who were
    (on medication when taking the test) vs (people who were not)"""
    df = pd.read_csv(DATA_SOURCE)

    MED = df.loc[df["currently_medicated"] == "Yes"]
    NON_MED = df.loc[df["currently_medicated"] == "No"]

    concatenated = pd.concat(
        [NON_MED.assign(dataset="No MED"), MED.assign(dataset="MED")]
    )
    sns.boxplot(y="currently_medicated", x="reaction_time", data=concatenated)

    plt.savefig('view_data_currently_medicated.png', dpi=1000)
    plt.show()
```



In [30]:

```
def get_data(df, field, ans):
    d = df.loc[df[field] == ans]
    print(f"{field} = {ans}", list(d["reaction_time"].to_numpy()), "\n")
```

In [31]:

```
def data_arrays():
    df = pd.read_csv(DATA_SOURCE)

    get_data(df, "currently_medicated", "Yes")
    get_data(df, "currently_medicated", "No")

    get_data(df, "adhd", "Yes")
    get_data(df, "adhd", "No")

data_arrays()
```

```
currently_medicated = Yes [0.428, 0.2865, 0.9226, 0.758, 0.2934]

currently_medicated = No [0.3496, 0.5686, 0.7577, 0.2084, 0.3046, 0.286, 0.2968, 0.289, 0.3592, 0.2894]

adhd = Yes [0.3496, 0.428, 0.5686, 0.7577, 0.2084, 0.2865, 0.3046, 0.286, 0.9226, 0.758, 0.2968, 0.289, 0.2934,
0.3592, 0.2894]

adhd = No [0.9222, 0.748, 0.6458, 0.6702, 0.4472, 0.7648, 0.321, 1.3816, 0.3796, 0.9236, 0.348, 0.62, 0.279999,
0.8388, 0.24, 0.3266, 0.2268, 0.2456, 0.732, 0.644, 0.687, 0.783, 0.3326, 1.402, 0.3228, 0.351, 0.307, 0.265, 0
.49, 0.3452, 0.271, 0.3174, 0.2216, 0.3338, 0.7528, 0.3576, 0.3304, 0.3136, 0.3694, 0.3216, 0.4, 0.2786]
```

In [32]:

```
def two_sample_t_test_adhd_vs_no_adhd():
    df = pd.read_csv(DATA_SOURCE)

    ADHD = df.loc[df["adhd"] == "Yes"]
    NON_ADHD = df.loc[df["adhd"] == "No"]

    ADHD = ADHD["reaction_time"]
    NON_ADHD = NON_ADHD["reaction_time"]

    return stats.ttest_ind(a=ADHD, b=NON_ADHD, equal_var=False)

two_sample_t_test_adhd_vs_no_adhd()
```

Out[32]:

```
Ttest_indResult(statistic=-1.105027385945976, pvalue=0.27721430739901304)
```

In [33]:

```
def two_sample_t_test_med_vs_no_med():
    df = pd.read_csv(DATA_SOURCE)

    ADHD = df.loc[df["currently_medicated"] == "Yes"]
    NON_ADHD = df.loc[df["currently_medicated"] == "No"]

    ADHD = ADHD["reaction_time"]
    NON_ADHD = NON_ADHD["reaction_time"]
```

```
    return stats.ttest_ind(a=ADHD, b=NON_ADHD, equal_var=False)


two_sample_t_test_med_vs_no_med()
```

Out[33]:
```
Ttest_indResult(statistic=1.200131979330287, pvalue=0.28039419522305686)
```

In [ ]:
```

```