研究影响Estimated time（Etime）的因素

变量：temporal relevance，satisfaction ，task difficulty，time-critical

1. 固定其他几个因素，研究High temporal relevance和Low temporal relevance的影响：
   1. 考虑Etime的值：

固定其他几个因素时，在High和Low条件下Etime值的大小上并不能发现具有一致性的规律。有的是High条件下的Etime值大，有的是Low条件下的Etime值大。（task11-14：easy&time-critical，task1-6,15：easy&non-time-critical，task7-10,16-20：difficult&non-time-critical）

* 1. 考虑Etime和Dtime的差值：

首先可以考虑差值的正负，这反映用户是将时间估计得更长还是更短，从正负来看，high和low没有很明显的区别，都是有一些正值一些负值，相对来说low的正负分布比较平均，high的负值要少一些。

其次可以考虑差值的大小，无论是原始值的大小还是绝对值的大小，high和low都没有很明显的区别。不过原始差值大体上high条件下比low条件下大，而绝对差值二者差别不大，从这一点来看原始差值比绝对差值可能更能反映时间感知的变化。

* 1. 考虑偏差率的大小，及差值与Dtime的比值的大小：

无论是原始差值的偏差率还是绝对差值的偏差率，high和low条件下依然没能发现很明显的区别。不过原始偏差率大体上high条件下比low条件下大，而绝对偏差率二者差别不大，从这一点来看原始偏差率比绝对偏差率可能更能反映时间感知的变化。

总的来看，固定其他几个因素，只研究temporal relevance对于Estimated time的影响，并不能直接发现high和low两种条件下有很明显的区别。不过对比发现用Etime与Dtime的原始差值或偏差率作为衡量时间感知的变量比较合适。

1. 固定其他几个因素，研究sat，midsat，unsat三种设置的影响：
   1. 考虑Etime的值：

Sat，midsat，unsat三种不同设置下，Etime的值的相对大小并没有明显的规律。

* 1. 考虑Etime和Dtime的差值：

首先考虑差值的正负，三种不同设置下差值都是有正有负，没有显著区别。

其次考虑差值的大小，对原始值和绝对值的大小都进行了分析，不过并不能发现三种不同设置下有显著的区别。

* 1. 考虑偏差率的大小：

无论是原始差值的偏差率还是绝对差值的偏差率，三种不同设置下依然没有发现显著的区别。

总的来看，固定其他几个因素，只研究satisfaction的设置对于Estimated time的影响，并不能直接发现sat，midsat和unsat三种不同设置下有显著的区别。一个可能的原因是我们对sat，midsat和unsat三种设置不太合理，即对搜索结果进行颠倒排序这种方法并不能使得sat，midsat和unsat三种设置达到我们预想的效果。

1. 固定其他几个因素，研究easy task和difficult task的影响（注意time-critical的task都是easy的task，所以在这里没有讨论time-critical的4个task）：

第3部分和第4部分在图中要关注的应该是整个两组数据之间的大小关系，而不只是对应点的大小关系，因为它们是不同的task

* 1. 考虑Etime的值：

对7个easy的task和9个difficult的task的值进行了分析，通过显著性检验没有发现easy和difficult的不同条件下有显著差异。

* 1. 考虑Etime和Dtime的差值：

首先考虑差值的正负，easy和difficult条件下没有显著差异。

其次考虑差值的大小，对原始值和绝对值的大小都进行了分析，也没有发现easy和difficult条件下有显著差异。

* 1. 考虑偏差率的大小：

无论是原始差值的偏差率还是绝对差值的偏差率，easy和difficult这两种不同条件下依然没有发现显著的差异。

总的来看，固定其他几个因素，只研究task difficulty对于Estimated time的影响，并不能直接发现easy和difficult两种不同条件下有显著的差异。一个可能的原因是我们对easy和difficult的标注并不十分合理，完全是根据我们主观判断的。

1. 固定其他几个因素，研究time-critical和non-time-critical的影响（注意difficult的task都是non-time-critical的task，所以在这里没有讨论difficult的9个task）：
   1. 考虑Etime的值：

对4个time-critical的task和7个non-time-critical的task的值进行了分析，通过显著性检验没有发现time-critical和non-time-critical的不同条件下有显著差异。

* 1. 考虑Etime和Dtime的差值：

首先考虑差值的正负，没有发现time-critical和non-time-critical条件下有显著差异。

其次考虑差值的大小，对原始值和绝对值的大小都进行了显著性分析，没有发现time-critical和non-time-critical条件下有显著差异。

* 1. 考虑偏差率的大小：

无论是原始差值的偏差率还是绝对差值的偏差率，easy和difficult这两种不同条件下依然没有发现显著的差异。

总的来看，固定其他几个因素，只研究time-critical对于Estimated time的影响，并没有发现time-critical和non-time-critical两种不同条件下有显著的差异。一个可能的原因是我们对time-critical的task的设置并不一定起到了作用。

通过对上面这些变量的分析（直接控制其他变量），总体上并没有找到如我们预期的那样值得注意的规律，我想到有两种可能的解释（不考虑我们实验数据太少…），也是我觉得可能继续分析的两个方向：

1. 这些因素对于时间感知的影响是相互交织在一起的，如果固定其他变量只研究某一个变量的影响很难找到一个一致性的规律，这也可能说明这些因素对于时间感知的影响不是简单的线性相加的，那么可以去进行多变量的研究。
2. 可能还有一些变量的影响我们不能忽略，必须考虑。我想到的有两个可以考虑的变量，一个是dwell time，当dwell time不同的时候用户估计时间产生的偏差是不一样的，比如dwell time比较小的时候用户可能估计得准一些，那么我们在研究Etime的时候应该把Dtime的值考虑进去；另一个是user，也就是说不同的user估计时间的能力是不一样的，那么是不是可以根据每个user的数据对他对应的Etime做个归一化之类的。