In the next iteration of this design I would consider which colors should be used. This particular set is useful for distinguishing the lighting groups, but looks a bit too much like Easter eggs. The map should also be slightly scaled down.

In each case I would suggest slight illumination of the switches and map themselves. This way, even in the dark, the user will be able to quickly glance and see which switch controls which lights.

Question 3: Formal Experimentation Design

- **Hypotheses**

Groups using the keypad input device will have faster data entry rates and lower error rates.

Data look up speeds for finding addresses will be quicker for the group using the stylus.

Color screens will not affect data entry rates or error rates, rates will be solely dependent on input device.

Groups who use the color screens will rate their satisfaction with the device higher than those with black and white screens.

- **Independent Variables, and their levels**

Input Device (Stylus, Keypad)

Monitor Type (Color, Black and White)

- **Dependent Variables, and their levels**

Error Rate (Number of incorrectly entered values per total number of values entered)

Data Entry Speed (Seconds required to enter all data in a given task divided by number of values entered)

Data Look-up Speed (Seconds required to find all entries in a given task divided by the number of values looked-up)

User Satisfaction (Subjective rating from 1-7 given by the user as to their satisfaction with device for data entry purposes)

- **Tasks that participants will do** (label them as A, B, C, etc)

Task A:
Users must read down a list of 60 seven-digit numbers and enter each one sequentially into the device. Users will be given a score of 1000 points and told that the value will drop by 1 every second that passes and increase by 1 for every correct value entered into the system. They will be asked to try to maximize their score. The users will be timed and number or errors noted by the system. These values will not be reported to the users during the test, only the score is shown.

Task B:
Users are given a list of 40 street addresses and asked to find and identify them within the system's address database. Again the users are given a starting score of 1000 which will tick down by 1 every second. Users will be timed but no errors reported, the system will only let the user advance after they have correctly found the corresponding address.

At this point an assumption has been made that the handheld device will contain a previously entered list of houses to be metered. It is possible that this information could be automatically collected (GPS), or that it needs to be entered by the reader each time she visits a house. For now we will assume a pre-existing data set that needs to be traversed before each reading. The actual implementation of this feature will greatly effect the need to study errors and speed of search. This will require a study of its own. Here I just hope to discover the effect of the input device and screen type on another activity for which the device might be used.

- **Blocking of experiment** (Including within-subjects versus between-subjects, how many times each task is done, which levels of each factor are used in each condition, how many participants are in each condition.)

Color Screen (16)		**Black & White (16)**	
Stylus (8)	Keypad (8)	Stylus (8)	Keypad (8)
A	B	B	A
B	A	A	B

The two display types will be run as a between group study. This way we can hope to gage a satisfaction rating of the system without users of the black and white system being biased by what might be seen as the flashier, color device. All participants will be given the satisfaction questionnaire at the end of their two tasks.

To avoid learning effects that could possibly result from task B, we will break down the input types again into a between groups test. This also requires participants to only run through each task once. The assumption is that the learning curve for stylus number entry and keypad entry is quite low. I have assumed that by 60 iterations average speed per entry will have leveled out.

Each of the participants within a input device and display subgroup will be asked to perform both tasks. A latin square was used to reduce learning effects through familiarity with the input device. In addition, since I hypothesized that display type would not effect error rate or input speed, it may also be possible to average input types together across display types.

Further experiments can be run to discover how data length affects input speed and error rate as well as address how housing address data should be retrieved or entered. In addition, we can test to see if speeds increase for input devices beyond 60 entries. For now we should focus on the device's primary application: the house-by-house entering of meter readings.