More on experimental units—The Sling Shot Problem Revisited—Detroit vs. Nascar

Suppose we are interested in quantifying the effects of stem design (U vs. V) and elastic weight (heavy, medium, light) on the distance traveled by a marble shot from any slingshot made with one of these 6 treatments. Distance travelled will be measured from a piece of tape behind which the shooter stands (with his toes on the tape), to the final resting place of the marble. The same person will shoot all marbles in the same manner.

The population(s) of interest:

Scenario 1:
Suppose we physically create 6 slingshots, one for each treatment. The two-letter abbreviations below represent a slingshot that you could pick up and use.

UH    UM    UL
VH    VM    VL

We make three “tosses” per treatment, with the exact same marble (pain in the butt, yes).

What is the experimental unit?

What is (are) the population(s) about which we can draw conclusions?

Does this experiment allow us to answer the question of interest?

Scenario 2:
Same as Scenario 1 except that we make three “tosses” per treatment, each toss with a different marble, but all marbles were purchased in such a way as to ensure they were of the same brand, size, and weight, with only manufacturing errors in size and weight. A total of 18 marbles were used.

What is the experimental unit?
What is (are) the population(s) about which we can draw conclusions?

Does this experiment allow us to answer the question of interest?

Scenario 3:
Suppose we physically create 18 slingshots, three for each treatment. The two-letter abbreviations below represent a slingshot that you could pick up and use.

UH    UH    UH    UM    UM    UM    UL    UL    UL
VH    VH    VH    VM    VM    VM    VL    VL    VL

We make three “tosses” per treatment, \textit{with the exact same marble} (pain in the butt, yes).

What is the experimental unit?

What is (are) the population(s) about which we can draw conclusions?

Does this experiment allow us to answer the question of interest?

Scenario 4:
Same as Scenario 3 except that we make three “tosses” per treatment, each toss with a different marble, but all marbles were purchased in such a way as to ensure they were of the same brand, size, and weight, with only manufacturing errors in size and weight. A total of 18 marbles were used.

What is the experimental unit?

What is (are) the population(s) about which we can draw conclusions?

Does this experiment allow us to answer the question of interest?

What are the pros and cons of Scenario 3 vs 4? Remember—think “real world,” not “correct answer.”