Dolly’s hypothesis “experiencing classical music leads to an enhancement in short-term memory” means not a correlation but a causal relation, which means she shouldn't be using a survey research method which is a type of descriptive research and should instead do an experimental research. Additionally, a rating of one’s memory ability by oneself reflects only one’s subjective opinion instead of the objective fact. By the two reasons above, I conclude that Dolly’s finding is not very evidential for her hypothesis.

As for an alternative study to test her hypothesis more rigorously, I propose an experimental research with the following operational definition:

(1) The test subjects are 20 undergraduate students who have no classical music lessons experience before.

(2) After obtaining informed consent, the subjects are randomly assigned to either the experimental group or the control group in half by a computer.

(3) The treatment for the experimental group is having free 1-hour classical music lessons twice a week for a period of 3 months. The control group is required to avoid any form of classical music experience that can exceed 5 minutes in duration.

(4) The assessment for subjects' short-term memory is the Forward Digit Test (Raviv, 2014), which the experimenter should follow the following guidelines (DIGIT SPAN (DGS), n.d.):  
 a. Before the assessment starts, the experimenter should instruct the subject that the experimenter will present a series of digits verbally and the subject’s task is to repeat the same series of digits verbally in the order that was presented.  
 b. The experimenter then should start with the section of 3-digit series on a randomly generated list of n-digit series by computer beforehand. If the subject responds correctly, then they will move on to a series of one-more-digit in span. If the subject fails on occasions less than 3, the experimenter should pick the next series of digits with the same length in span and redo the presentation. If the subject fails on occasions of 3 or more, the assessment terminates.

c. The score of the subject on the assessment should be proportional to the longest span reached successfully and disproportional to the number of unsuccessful attempts.

(5) The independent variable is the total hours of classical music lessons and the dependent variable is the score on the Forward Digit Test. The assessment should be conducted before, after 1 month, after 2 months, and after 3 months of treatment to track progress of subjects over time.

Reference

Raviv, O. (https://psychology.stackexchange.com/users/38/ofri-raviv) What test could I use

to measure the short-term memory capacity of a group of people? (2014, May 10).

Psychology & Neuroscience Stack Exchange.

https://psychology.stackexchange.com/a/6412/31815

DIGIT SPAN (DGS). (n.d.). Cambridge Cognition. Retrieved November 3, 2022, from

https://www.cambridgecognition.com/cantab/cognitive-tests/memory/digit-span-dgs/