***Appendix 3***

**Study 5**

Study 5 explored whether different steps are implicitly associated with different difficulties (easy vs. hard) in human minds.

**Method**

**Participants**

One hundred and thirty-two participants (105 females; mean age = 20, *SD* =1.33) recruited via social media groups from our university completed the study in a laboratory. They answered the questions on an online survey platform ([www.wjx.cn](http://www.wjx.cn)). They got payment for their participation.

**Materials and procedure**

This study adopted a within-participant design. It used an implicit association paradigm to assess participants' perception of difficulty about Step 1 and Step 5. The participants were instructed to make classification as fast as possible, while make as few mistakes as possible*.*

The task included 7 stages as described below and diagrammed in Figure 5.

In Stage 1 (step classification), participants needed to classify steps as the first or final step. The purpose was to prepare the participants for later joint tasks. At the bottom of the screen, a diagram showed five steps of a scenario by repeatedly presenting a picture resembling that scenario five times. At the top of the screen was shown the categories of "the first step" and "the final step". When the first (final) step of the diagram was circled, the participants needed to choose the category "the first (final) step". In addition, "the first step" ("the final step") always appeared on the left (right) of the top part of the screen whenever they appeared in any stage. Six scenarios appearing in Study 3 were used: a lottery game, a knowledge contest, guessing the names of songs, a certificate examination, an electronic game, and a water game. Each scenario appeared twice: once with circle on the first icon, and the other with circle on the final icon. Totally, this stage had 12 trials.

In Stage 2 (difficulty classification – easy on the left), participants needed to classify activities as easy or hard. The purpose was to prepare the participants for later joint tasks. At the bottom of the screen was shown an activity. At the top of the screen was shown the categories of "easy" and "hard". When the activity was easy (like primary math) or hard (like advanced math), the participants needed to choose the category "easy" or "hard" respectively. Six easy topics were used: primary school mathematics, jogging 50m, primary school composition, watching films, reading books, and simple drawing. Six hard topics were used: advanced mathematics, long run 5km, doctoral dissertation, making film, writing a book, and character sketches. Totally, this stage contained 12 trials.

Then came Stage 3 (the joint compatible task - unpracticed). In this stage, when the bottom of the screen showed a scenario, the participants needed to classify the circled step as the first or final step; when it showed an activity, they needed to classify the activity as easy or hard. At the top of the screen, the category options paired the first step with easy (on the left) and the final step with hard (on the right) together. The bottom of the screen copied that of Stages 1 and 2. As a result, this stage had 24 trials.

Stage 4 (the joint compatible task - practiced) completely replicated Stage 3. We assumed that the participants were relatively unpracticed and practiced in Stages 3 and 4 respectively.

Stage 5 (difficulty classification – hard on the left) replicated Stage 2 except that "hard" ("easy") was on the left (right) now. This is a preparation for the next stage, which would set "hard" on the left.

Stage 6 (the joint incompatible task - unpracticed) replicated Stage 3 except that at the top of the screen, the category options paired the first step with hard (on the left) and the final step with easy (on the right) together.

Stage 7 (the joint incompatible task - practiced) completely replicated Stage 6. We assumed that the participants were relatively unpracticed and practiced in Stages 6 and 7 respectively.

*Half participants did the above version of 7 stages, having the compatible task before the incompatible task. The other half participants did a counterbalanced version: stages 2 to 4 were swapped with stages 5 to 7, i.e., they had the incompatible task before the compatible task.*

In any trial of any stage, if a participant made a wrong classification, the procedure would ask him/her to redo that trial until he/she got it correct. For each trial, only the response time of the correct response was recorded and analyzed.

If there was no implicit association between steps and difficulty, then the response time in the joint compatible unpracticed task and joint incompatible unpracticed task should be equal; otherwise, unequal. So should the response time in the joint compatible practiced task and joint incompatible practiced task.

the first step

the final step



easy

primary mathematics

hard

the first step

or

easy



the final step

or

hard

the first step

or

easy

the final step

or

hard

or

primary mathematics

the first step

or

easy



the final step

or

hard

the first step

or

easy

the final step

or

hard

or

primary mathematics

hard

primary mathematics

easy

the first step

or

hard



the final step

or

easy

the first step

or

hard

the final step

or

easy

or

primary mathematics

the first step

or

hard



the final step

or

easy

the first step

or

hard

the final step

or

easy

or

primary mathematics

Figure 5. The seven stages of the implicit association test. Each row of screen(s) represents a stage. Two screens in one row mean, in that stage, participants would encounter two types of screens (questions). Each downward arrow indicates going from the last stage to the next stage.

**Results**

We firstly calculated each participant's average reaction time during the *joint unpracticed* compatible and incompatible tasks (stages 3 and 6) respectively. We then calculated the overall statistics as follows. The average reaction time was 910 (SD = 258) milliseconds for the compatible task and 1196 (SD = 422) milliseconds for the incompatible task. Paired-sample statistics showed that *95% confidence interval* = [-354, -216], *Cohen’s d* = -0.712.

For the practiced conditions (stages 4 and 7), we firstly calculated each participant's average reaction times during the *joint practiced* compatible and incompatible tasks respectively. We then calculated the overall statistics as follows. The average reaction time was 759 (*SD* = 217) milliseconds for the compatible task and 896 (*SD* = 256) milliseconds for the incompatible task. Paired-sample statistics showed that *95% confidence interval* = [-173, -103], *Cohen’s d* = -0.679.

In other words, people reacted faster in the compatible task than in the incompatible task in both unpracticed and practiced situations, which suggested that there was an implicit association between the first step and easy and the final step and hard. That is, people had a stereotype that the first step is easy and the final step is hard. Studies 1 to 3 all told participants that each step had equal difficulty and pass rate, but the stereotype that the final step is hard might unconsciously seduced them to devote the most valuable resources into the final step. This is consistent with previous finding that stereotype can have a major impact on decision making, as is revealed in the "representativeness heuristic" (Erbas & Ocal, 2022; Galesic et al., 2018; Krawczyk & Rachubik, 2019; Tversky & Kahneman, 1974; Walston et al., 2022).

Erbas, A. K., & Ocal, M. F. (2022). Students' intuitively-based (mis)conceptions in probability and teachers' awareness of them: the case of heuristics [Article; Early Access]. *International Journal of Mathematical Education in Science and Technology*. <https://doi.org/10.1080/0020739x.2022.2128454>

Galesic, M., Goode, A. W., Wallsten, T. S., & Norman, K. L. (2018). Using Tversky's contrast model to investigate how features of similarity affect judgments of likelihood [Article]. *Judgment and Decision Making*, *13*(2), 163-169. <Go to ISI>://WOS:000428850500002

Krawczyk, M. W., & Rachubik, J. (2019). The representativeness heuristic and the choice of lottery tickets: A field experiment [Article]. *Judgment and Decision Making*, *14*(1), 51-57. <Go to ISI>://WOS:000457283300005

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty - heuristics and biases [Article]. *Science*, *185*(4157), 1124-1131. <https://doi.org/10.1126/science.185.4157.1124>

Walston, Z., Whelehan, D. F., & O'Shea, N. (2022). Clinical decision making in physical therapy - Exploring the 'heuristic' in clinical practice [Article]. *Musculoskeletal Science and Practice*, *62*, Article 102674. <https://doi.org/10.1016/j.msksp.2022.102674>