

ESM 1: Comments on the Experimental Multimatrix Design

We implemented a rotated multimatrix design in the present experiment, in order to achieve the following three important goals:

- (1) Applying a high number of test items (48 parallel items = 96 items in total) in order to gain reliable results and effect size estimates without putting excessive demands on the participating schoolchildren.
- (2) Realizing a within-subject design according to the multimedia manipulation of items to allow for a separate estimation of test scores for text-only and text-picture items, relying on an IRT analysis. This was especially important in order to assure a useful investigation of interactions between several student characteristics and their influence on students' performance in text-only compared to text-picture items.
- (3) Balancing item positions for text-only and text-picture items to prevent from potential bias due to students exhaustion or other test-length effects which might interact with the experimental manipulation.

In compliance with this, we arranged four defined item blocks (12 items each) in such a way that every item that was solved by a certain student in a certain mode (text-only or text-picture) was linked to the remaining items that were not presented in the booklet, as well as to the parallel items that were, of course, not solved again by the same student in the opposite mode (text-only/text-picture). This was achieved by always combining three test blocks according to a well-established principle:

	PART 1	PART 2	PART 3
Booklet 1	BLOCK_1 (TO)	BLOCK_2 (TP)	BLOCK_4 (TP)
Booklet 2	BLOCK_2 (TO)	BLOCK_3 (TP)	BLOCK_1 (TP)
Booklet 3	BLOCK_3 (TO)	BLOCK_4 (TP)	BLOCK_2 (TP)
Booklet 4	BLOCK_4 (TO)	BLOCK_1 (TP)	BLOCK_3 (TP)
Booklet 5	BLOCK_1 (TP)	BLOCK_3 (TO)	BLOCK_2 (TO)
Booklet 6	BLOCK_2 (TP)	BLOCK_4 (TO)	BLOCK_1 (TO)
Booklet 7	BLOCK_3 (TP)	BLOCK_1 (TO)	BLOCK_4 (TO)
Booklet 8	BLOCK_4 (TP)	BLOCK_2 (TO)	BLOCK_3 (TO)

Figure 1. Visualization of the assignment of four experimentally manipulated item blocks (*text-only* = TO vs. *text-picture* = TP) to eight booklets following a rotated multimatrix design.

REPRESENTATIONAL PICTURES IN ASSESSMENT

The design in Figure 1 represents identical item sets in different shades and with corresponding labels (e.g., BLOCK_1 = Block 1, representing 12 items), while every test block was either presented as a text-only (TO) or as a text-picture (TP) version, explicated by the index in brackets.

Thus, every row displays one of eight booklets with a certain constellation of 36 items (including at least one text-only *and* one text-picture block) that were randomly assigned to the participating students. Furthermore, to balance the item positions, every test item was presented as a text-picture and as a text-only item once in all three test parts. As mentioned in the manuscript, booklets were shortened to 24 items (two blocks) for the regional school, because of a time restriction. This was accounted for in the design by eliminating the third part (last column) of each booklet, which did not harm the linking between the items or the balancing of item positions at all.