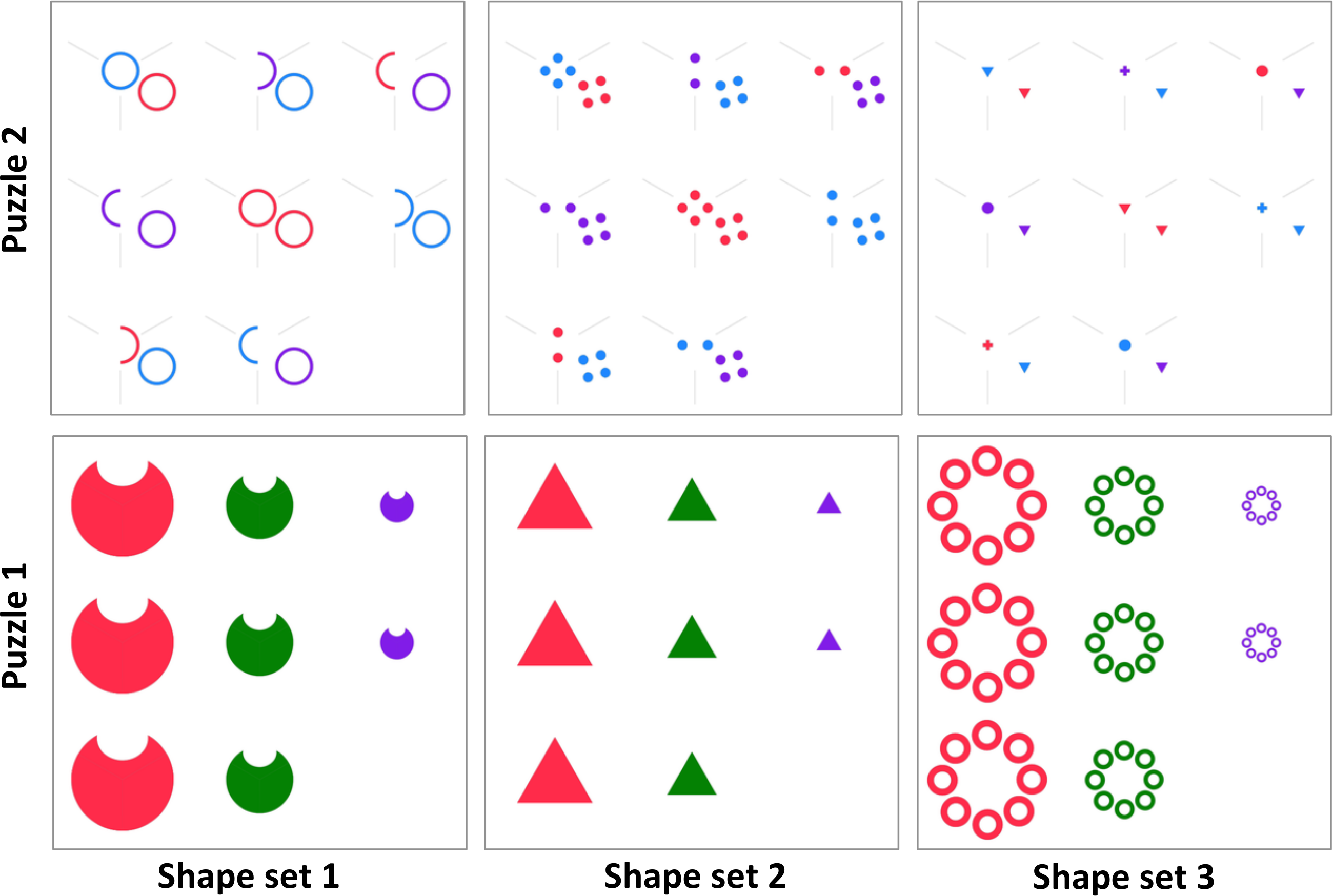
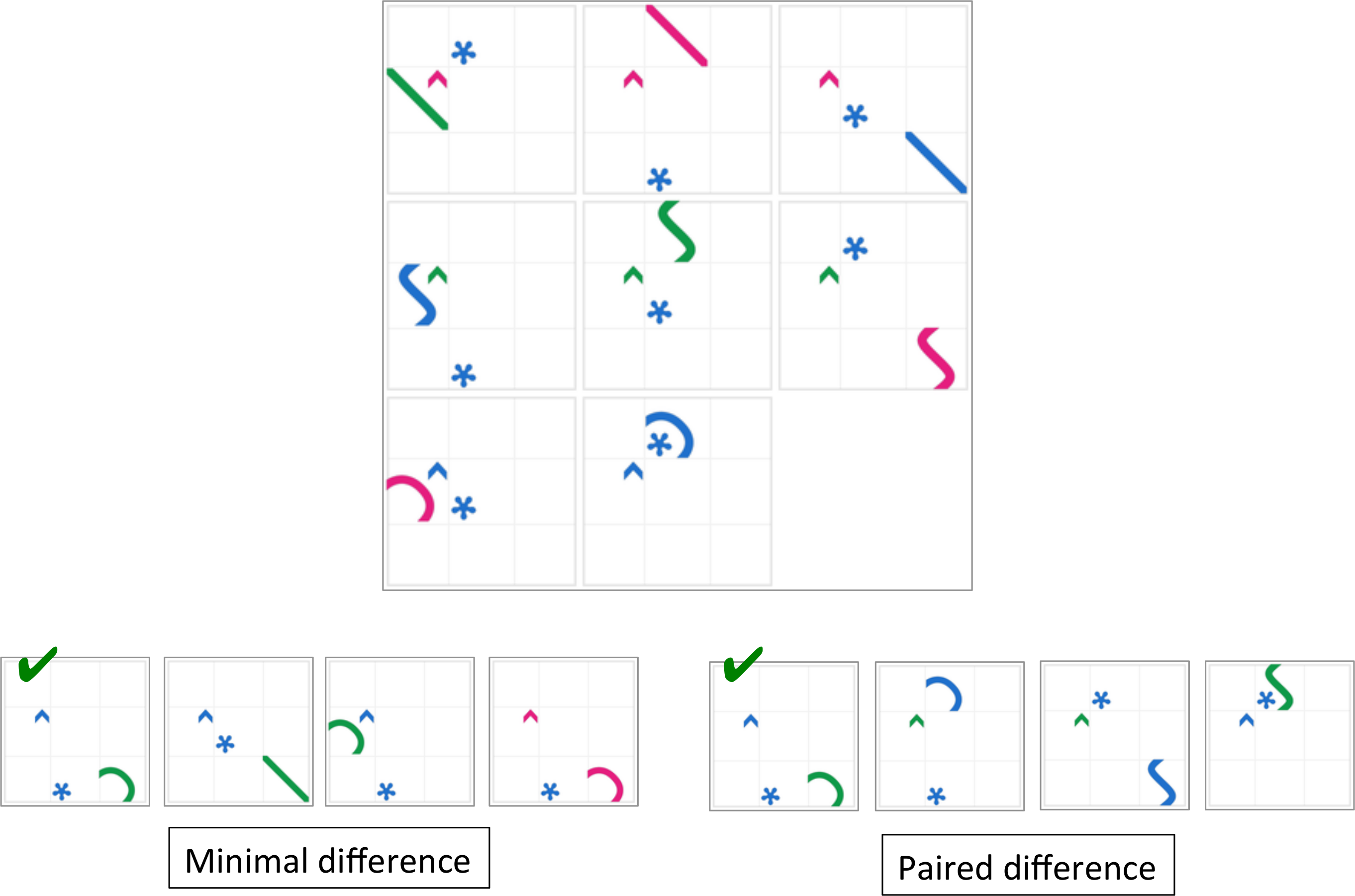
**Supplementary Material**

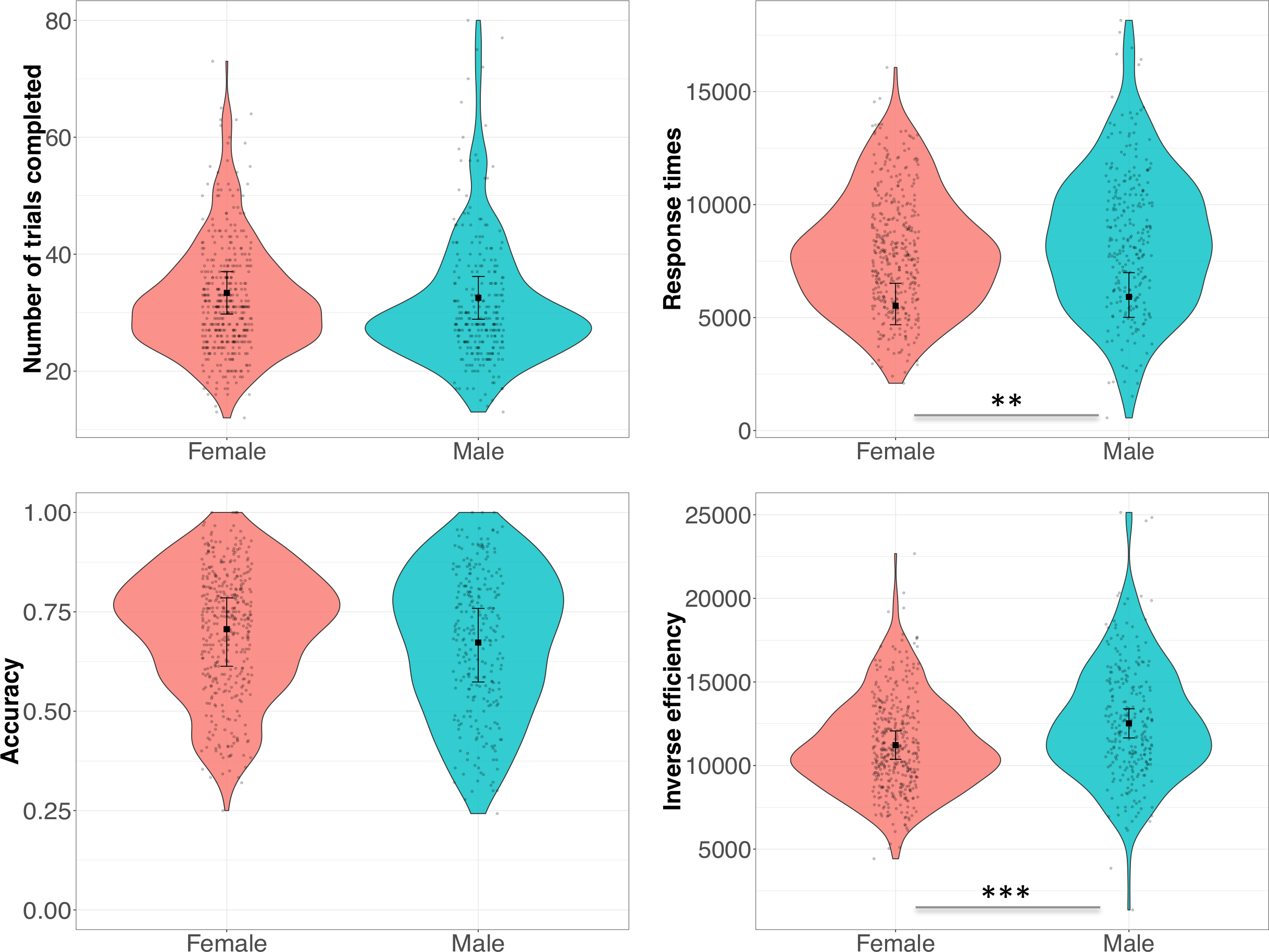
**Supplementary Figures**

****

*Figure S1.**Shape Sets*. Examples of the three different shape sets for two puzzles.



*Figure S2.**Target and Distractor Generation.* The minimal difference strategy potentially allows solving the puzzle by looking at the solutions only (e.g., one of the solutions is “minimally different” from all the others). The paired difference strategy does not allow this but potentially allows “pop-out” effects. Green tick marks indicate the correct response option.



*Figure S3.**Gender Differences in MaRs-IB Performance.* Figure shows the number of items completed, accuracy, response times and inverse efficiency. \*\* *p* < .01, \*\*\* *p* < .001.

**Supplementary Tables**

*Table S1. Descriptive Statistics of the MaRs-IB by Test Form.*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **puzzle set 1** | **puzzle set 2** | **puzzle set 3** |
| Median RT | 8372.50 | 7530.75 | 7790.00 |
| IQR RT | 3606.0 | 4051.5 | 4338.5 |
| Mean correct (%) | 69.45 | 69.04 | 69.01 |
| SE correct (%) | 1.09 | 1.18 | 1.11 |
| Min correct (%) | 27.78 | 25.00 | 24.24 |
| Max correct (%) | 100 | 100 | 100 |
| Mean N items | 31.29 | 32.32 | 32.19 |
| SE N items | 0.70 | 0.61 | 0.74 |
| Min N items | 13 | 12 | 14 |
| Max N items | 77 | 64 | 80 |

*Table S2. Descriptive Statistics of the MaRs-IB by Shape Set.*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **shape set 1** | **shape set 2** | **shape set 3** |
| Median RT | 7713.50 | 8198.75 | 7855.25 |
| IQR RT | 4548.50 | 5234.62 | 5325.25 |
| Mean correct (%) | 69.83 | 69.10 | 68.08 |
| SE correct (%) | 0.79 | 0.80 | 0.80 |
| Min correct (%) | 10.00 | 9.09 | 18.18 |
| Max correct (%) | 100 | 100 | 100 |
| Mean N items | 10.65 | 10.66 | 10.63 |
| SE N items | 0.13 | 0.13 | 0.14 |
| Min N items | 4 | 4 | 3 |
| Max N items | 27 | 27 | 26 |

*Table S3. Descriptive Statistics of the MaRs-IB by Distractor Strategy.*

|  |  |  |
| --- | --- | --- |
|  | **minimal difference** | **paired difference** |
| Median RT | 7853.00 | 7968.75 |
| IQR RT | 4344.38 | 4781.00 |
| Mean correct (%) | 68.91 | 69.45 |
| SE correct (%) | 0.71 | 0.73 |
| Min correct (%) | 18.75 | 18.18 |
| Max correct (%) | 100 | 100 |
| Mean N items | 15.96 | 15.98 |
| SE N items | 0.2 | 0.2 |
| Min N items | 5 | 5 |
| Max N items | 40 | 40 |

*Table S4. Descriptive Statistics of the MaRs-IB by Age Group and Gender.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **statistic** | **younger adolescents** | **mid-adolescents** | **older adolescents** | **adults** |
|  | *11.27 - 13.40y* | *13.40 - 15.92y* | *15.93 - 17.99y* | *18.00 - 33.15y* |
| Median RT | 6944 (m=7005.5 f=6871.75) | 7552.5 (m=7964.5 f=6841) | 7951.75 (m=8774.25 f=7582) | 9454.5 (m=10276 f=9304) |
| IQR RT | 3879 (m=4301.25 f=3771.5) | 3722.75 (m=4629.25 f=2669) | 3309.25 (m=3337.88 f=2780.88) | 3230.88 (m=3011.5 f=3297) |
| Mean correct (%) | 60.7 (m=57.37 f=62.58) | 67.52 (m=67.49 f=67.56) | 72.56 (m=72.11 f=72.87) | 80.82 (m=78.85 f=81.43) |
| SE correct (%) | 1.25  (m=2.22 f=1.48) | 1.18  (m=1.79 f=1.54) | 1  (m=1.69 f=1.23) | 1.3  (m=3.27 f=1.37) |
| Mean N items completed | 33.68 (m=34.4 f=33.27) | 32.29 (m=32.17 f=32.42) | 28.99 (m=28.24 f=29.52) | 33.43 (m=31.96 f=33.89) |
| SE N items completed | 0.74  (m=1.26 f=0.92) | 0.76  (m=1.19 f=0.92) | 0.75  (m=1.22 f=0.94) | 0.93  (m=2.07 f=1.04) |

*Note.* m = males, f = females.

*Table S5. Accuracy: Age Group Contrasts.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***contrast*** | ***estimate*** | ***SE*** | ***p*Bonf.** | ***p*uncorr.** |
| Younger adolescents - Mid adolescents | -0.28 | 0.13 | .163 | .027 |
| Younger adolescents - Older adolescents | -0.33 | 0.2 | .592 | .099 |
| Younger adolescents - Adults | -1.13 | 0.4 | .027 | .005 |
| Mid adolescents - Older adolescents | -0.06 | 0.16 | 1 | .728 |
| Mid adolescents - Adults | -0.85 | 0.39 | .159 | .026 |
| Older adolescents - Adults | -0.8 | 0.38 | .22 | .037 |

*Note.* Results are shown on the log odds scale.

*Table S6. Response Times: Age group Contrasts.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***contrast*** | ***estimate*** | ***SE*** | ***p*Bonf.** | ***p*uncorr.** |
| Younger adolescents - Mid adolescents | 0.03 | 0.04 | 1 | .542 |
| Younger adolescents - Older adolescents | 0.01 | 0.06 | 1 | .873 |
| Younger adolescents - Adults | -0.11 | 0.12 | 1 | .357 |
| Mid adolescents - Older adolescents | -0.01 | 0.05 | 1 | .783 |
| Mid adolescents - Adults | -0.14 | 0.12 | 1 | .253 |
| Older adolescents - Adults | -0.12 | 0.12 | 1 | .317 |

*Note.* Results are shown on the log scale.

*Table S7. Number of Items Completed: Age Group Contrasts.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***contrast*** | ***estimate*** | ***SE*** | ***p*Bonf.** | ***p*uncorr.** |
| Younger adolescents - Mid adolescents | 0.93 | 1.31 | 1 | .478 |
| Younger adolescents - Older adolescents | 0.12 | 2.02 | 1 | .951 |
| Younger adolescents - Adults | 1.51 | 5.17 | 1 | .772 |
| Mid adolescents - Older adolescents | -0.81 | 1.66 | 1 | .625 |
| Mid adolescents - Adults | 0.58 | 5.11 | 1 | .91 |
| Older adolescents - Adults | 1.39 | 5.11 | 1 | .788 |

*Table S8. Inverse Efficiency: Age Group Contrasts.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***contrast*** | ***estimate*** | ***SE*** | ***p*Bonf.** | ***p*uncorr.** |
| Younger adolescents - Mid adolescents | 762.94 | 413.23 | .395 | .066 |
| Younger adolescents - Older adolescents | 844.14 | 612.87 | 1 | .171 |
| Younger adolescents - Adults | 681.36 | 1317.95 | 1 | .612 |
| Mid adolescents - Older adolescents | 81.2 | 515.11 | 1 | .875 |
| Mid adolescents - Adults | -81.57 | 1300.44 | 1 | .951 |
| Older adolescents - Adults | -162.77 | 1318.15 | 1 | .903 |

*Table S9. MaRs-IB item statistics. N = number of responses; Dim = dimensionality; RT = median response time of correct responses (with interquartile range); IES = inverse efficiency*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Order** | **N** | **Dim** | **P-value (SE)** | **RT (IQR)** | **IES** |
| 1 | 648 | 2 | 97 (1) | 2412 (1579) | 2489 |
| 2 | 654 | 1 | 97 (1) | 4447 (2513) | 4566 |
| 3 | 646 | 1 | 98 (1) | 2464 (1188) | 2515 |
| 4 | 653 | 1 | 99 (0) | 3763 (1710) | 3786 |
| 5 | 647 | 1 | 100 (0) | 2528 (1381) | 2531 |
| 6 | 642 | 2 | 78 (2) | 9428 (7129) | 12154 |
| 7 | 644 | 1 | 98 (1) | 5184 (2923) | 5283 |
| 8 | 643 | 1 | 85 (1) | 7655 (5596) | 9031 |
| 9 | 653 | 1 | 94 (1) | 5899 (3017) | 6284 |
| 10 | 630 | 3 | 53 (2) | 15262 (10196) | 28787 |
| 11 | 628 | 3 | 59 (2) | 13966 (9770) | 23641 |
| 12 | 622 | 4 | 25 (2) | 16342 (12690) | 64332 |
| 13 | 619 | 3 | 49 (2) | 17343 (10754) | 35198 |
| 14 | 607 | 5 | 37 (2) | 18898 (12920) | 50983 |
| 15 | 641 | 3 | 56 (2) | 13600 (11373) | 24216 |
| 16 | 647 | 4 | 89 (1) | 7707 (6052) | 8642 |
| 17 | 638 | 4 | 59 (2) | 13944 (10089) | 23534 |
| 18 | 626 | 3 | 41 (2) | 15692 (11098) | 38223 |
| 19 | 635 | 2 | 82 (2) | 7902 (5100) | 9631 |
| 20 | 629 | 2 | 67 (2) | 12114 (7182) | 18099 |
| 21 | 596 | 7 | 49 (2) | 16676 (9658) | 33691 |
| 22 | 596 | 2 | 82 (2) | 7713 (4963) | 9362 |
| 23 | 591 | 4 | 81 (2) | 9847 (6226) | 12226 |
| 24 | 553 | 7 | 41 (2) | 15723 (11576) | 38303 |
| 25 | 542 | 2 | 77 (2) | 9358 (7050) | 12135 |
| 26 | 509 | 5 | 36 (2) | 17140 (10786) | 47674 |
| 27 | 477 | 3 | 54 (2) | 11507 (8290) | 21275 |
| 28 | 438 | 2 | 57 (2) | 13086 (7615) | 22928 |
| 29 | 392 | 7 | 62 (2) | 11622 (8504) | 18748 |
| 30 | 348 | 7 | 60 (3) | 9362 (5374) | 15663 |
| 31 | 316 | 3 | 55 (3) | 11404 (7901) | 20711 |
| 32 | 288 | 1 | 56 (3) | 8090 (6122) | 14472 |
| 33 | 253 | 1 | 60 (3) | 6414 (3698) | 10675 |
| 34 | 233 | 3 | 37 (3) | 8640 (8548) | 23139 |
| 35 | 210 | 8 | 42 (3) | 8321 (8830) | 19857 |
| 36 | 191 | 6 | 36 (3) | 6836 (8263) | 19203 |
| 37 | 176 | 3 | 36 (4) | 8496 (9960) | 23735 |
| 38 | 156 | 1 | 83 (3) | 4240 (2509) | 5087 |
| 39 | 135 | 2 | 41 (4) | 5448 (7342) | 13372 |
| 40 | 122 | 4 | 44 (5) | 5020 (4839) | 11341 |
| 41 | 113 | 1 | 84 (3) | 4248 (2618) | 5053 |
| 42 | 106 | 4 | 27 (4) | 5674 (6414) | 20739 |
| 43 | 91 | 1 | 33 (5) | 4720 (4510) | 14319 |
| 44 | 83 | 6 | 19 (4) | 2424 (4126) | 12574 |
| 45 | 73 | 7 | 33 (6) | 6316 (7412) | 19210 |
| 46 | 61 | 6 | 38 (6) | 4185 (5728) | 11099 |
| 47 | 57 | 2 | 21 (5) | 5035 (7364) | 23916 |
| 48 | 49 | 1 | 33 (7) | 2942 (6569) | 9010 |
| 49 | 47 | 2 | 47 (7) | 4181 (2674) | 8932 |
| 50 | 46 | 2 | 37 (7) | 4021 (4195) | 10880 |
| 51 | 41 | 2 | 29 (7) | 5712 (3462) | 19516 |
| 52 | 35 | 5 | 37 (8) | 1757 (2072) | 4730 |
| 53 | 30 | 4 | 20 (7) | 1286 (1058) | 6432 |
| 54 | 30 | 6 | 13 (6) | 1612 (1382) | 12094 |
| 55 | 28 | 6 | 36 (9) | 2770 (3133) | 7755 |
| 56 | 25 | 3 | 20 (8) | 1726 (1443) | 8630 |
| 57 | 21 | 1 | 43 (11) | 4151 (2261) | 9686 |
| 58 | 21 | 3 | 43 (11) | 3836 (2070) | 8951 |
| 59 | 20 | 5 | 25 (10) | 2806 (3091) | 11224 |
| 60 | 14 | 4 | 50 (14) | 1944 (1010) | 3888 |
| 61 | 16 | 2 | 12 (9) | 1590 (1312) | 12724 |
| 62 | 15 | 3 | 27 (12) | 5922 (4046) | 22208 |
| 63 | 14 | 5 | 14 (10) | 4792 (3122) | 33540 |
| 64 | 11 | 6 | 36 (15) | 1010 (452) | 2779 |
| 65 | 9 | 2 | 11 (11) | 849 (0) | 7641 |
| 66 | 9 | 8 | 22 (15) | 1194 (636) | 5375 |
| 67 | 8 | 5 | 25 (16) | 768 (59) | 3072 |
| 68 | 6 | 1 | 100 (0) | 1964 (527) | 1964 |
| 69 | 7 | 3 | 57 (20) | 3776 (1369) | 6608 |
| 70 | 8 | 2 | 50 (19) | 3495 (1420) | 6990 |
| 71 | 6 | 2 | 50 (22) | 4439 (2121) | 8878 |
| 72 | 5 | 4 | 40 (24) | 1164 (689) | 2910 |
| 73 | 6 | 4 | 17 (17) | 1033 (0) | 6198 |
| 74 | 4 | 5 | 50 (29) | 4274 (3827) | 8548 |
| 75 | 6 | 6 | 50 (22) | 989 (524) | 1978 |
| 76 | 4 | 5 | 75 (25) | 507 (327) | 676 |
| 77 | 3 | 3 | 33 (33) | 455 (0) | 1365 |
| 78 | 4 | 7 | 25 (25) | 631 (0) | 2524 |
| 79 | 5 | 4 | 0 (0) |  |  |
| 80 | 3 | 4 | 33 (33) | 640 (0) | 1920 |

*Table S10. MaRs-IB item statistics on items with no variance in completion rate. N = number of responses; Dim = dimensionality; RT = median response time of correct responses (with interquartile range); IES = inverse efficiency; ITC = item-total correlation. None of the items displayed uniform or non-uniform differential item functioning relative to age and gender.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Item** | **N** | **Dim** | **P-value (SE)** | **RT (IQR)** | **IES** | **ITC** |
| 6 | 348 | 2 | 75 (2) | 8064 (5357) | 10752 | 0.3 |
| 7 | 348 | 1 | 97 (1) | 4680 (2239) | 4819 | 0.13 |
| 8 | 348 | 1 | 81 (2) | 6522 (4959) | 8077 | 0.4 |
| 9 | 348 | 1 | 90 (2) | 5516 (2825) | 6114 | 0.31 |
| 10 | 348 | 3 | 50 (3) | 13642 (7592) | 27128 | 0.4 |
| 11 | 348 | 3 | 56 (3) | 11396 (7299) | 20442 | 0.25 |
| 12 | 348 | 4 | 19 (2) | 12684 (7888) | 66879 | 0.26 |
| 13 | 348 | 3 | 44 (3) | 14850 (9108) | 33557 | 0.22 |
| 14 | 348 | 5 | 33 (3) | 14443 (10586) | 43706 | 0.28 |
| 15 | 348 | 3 | 49 (3) | 11278 (7295) | 22819 | 0.31 |
| 16 | 348 | 4 | 86 (2) | 6528 (4228) | 7572 | 0.3 |
| 17 | 348 | 4 | 55 (3) | 11762 (8817) | 21543 | 0.19 |
| 18 | 348 | 3 | 31 (2) | 12776 (8538) | 41552 | 0.29 |
| 19 | 348 | 2 | 78 (2) | 6914 (4294) | 8846 | 0.35 |
| 20 | 348 | 2 | 56 (3) | 9958 (6473) | 17680 | 0.47 |
| 21 | 348 | 7 | 44 (3) | 13687 (8355) | 31336 | 0.42 |
| 22 | 348 | 2 | 79 (2) | 6600 (4052) | 8352 | 0.34 |
| 23 | 348 | 4 | 77 (2) | 8692 (4909) | 11329 | 0.4 |
| 24 | 348 | 7 | 34 (3) | 12947 (10951) | 37546 | 0.21 |
| 25 | 348 | 2 | 72 (2) | 8353 (4843) | 11674 | 0.36 |
| 26 | 349 | 5 | 34 (3) | 14413 (11665) | 42270 | 0.17 |
| 27 | 348 | 3 | 51 (3) | 10591 (6562) | 20706 | 0.37 |
| 28 | 348 | 2 | 55 (3) | 12398 (7486) | 22709 | 0.33 |
| 29 | 348 | 7 | 60 (3) | 10856 (8430) | 18164 | 0.4 |
| 30 | 348 | 7 | 60 (3) | 9362 (5374) | 15663 | 0.43 |

**Supplementary Information**

*Supplementary information SI1. Convergent validity (N = 50)*

In the manuscript, we report of a follow-up study that was conducted to assess convergent validity of the MaRs-IB. A power analysis suggested that 38 participants are sufficient to detect a correlation of 0.5 at 90% power. For this study, we originally recruited 50 participants, which were increased to 100 upon reviewer request. Below we report the results of the first sample (N = 50) (36 females, 14 males, mean age = 24.18, *SE* = 0.49, age range 20-35 years) for transparency on multiple rounds of data collection.

To assess convergent validity we inspected the product-moment correlation between MaRs-IB performance (at the aggregate level) and matrix reasoning scores in the ICAR (Condon & Revelle, 2014). The correlation was 0.54 [*t* (48) = 4.44, p < 0.001, 95% CI = 0.31 0.71], acceptable for convergent validity purposes (Carlson & Herdman, 2010). Linear regression also showed that correlation did not interact with the order in which participants did the two tasks (p = 0.716). The performance standard deviations of the ICAR matrix reasoning items in our sample and in Condon & Revelle (2014) were 0.25 and 0.5, respectively, warranting a correction for age restriction (Condon & Revelle, 2014). The range corrected correlation was 0.79, thus again within acceptable range. To assess the extent of divergent validity, we further compared the correlation described above to the correlations between the MaRs-IB and the remaining tasks of the ICAR (Table S11), namely, letter-number series completion, verbal reasoning and 3D rotations. The highest correlation with MaRs-IB performance was indeed with the ICAR matrix reasoning (Table 2). However, comparing the correlations using the Fisher r-z transform revealed that the MaRs-IB-ICAR matrices correlation did not significantly differ from the others (all ps > 0.24). Taken together, this evidence suggests an acceptable degree of convergent validity between the MaRs-IB and ICAR-matrix reasoning, with correlation sizes that are similar to those observed between the ICAR and other cognitive ability tests, such as the Shipley 2 (Condon & Revelle, 2014). We observe no evidence of divergent validity from other IQ-related measures of the ICAR, suggesting that the MaRs-IB may tap into a broad cognitive functioning construct.

*Table S11. Correlations between MaRs-IB and tasks from the International Cognitive Ability Resource (“MR” = matrix reasoning, “R3D” = 3D rotations, “LN” = letter and number series completion, “VR” = verbal reasoning).* \*\*\* p < 0.001, \*\* p < 0.01, \*p < 0.05, °p<0.1. Bonferroni corrected.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **MaRs-IB** | **MR** | **R3D** | **LN** |
| **MaRs-IB** |  |  |  |  |
| **MR** | 0.54\*\*\* |  |  |  |
| **R3D** | 0.43\* | 0.29 |  |  |
| **LN** | 0.39° | 0.46\*\* | 0.35 |  |
| **VR** | 0.45\*\* | 0.52\*\* | 0.38° | 0.37° |