Complexity and compositionality in fluid intelligence

John Duncanᵃᵇᶜ, Daphne Chylinskiᵃ, Daniel J. Mitchellᵇ, and Apoorva Bhandariᶜ

ᵃMedical Research Council, Cognition and Brain Sciences Unit, Cambridge CB2 7EF, United Kingdom;ᵇDepartment of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom; andᶜDepartment of Cognitive, Linguistic and Psychological Sciences, Brown University, Providence, RI 02912

Edited by Michael I. Posner, University of Oregon, Eugene, OR, and approved March 27, 2017 (received for review December 22, 2016)

Compositionality, or the ability to build complex cognitive structures from simple parts, is fundamental to the power of the human mind. Here we relate this principle to the psychometric concept of fluid intelligence, traditionally measured with tests of complex reasoning. Following the principle of compositionality, we propose that the critical function of fluid intelligence is splitting a complex whole into simple, separately attended parts. To test this proposal, we modify traditional matrix reasoning problems to minimize requirements on information integration, working memory, and processing speed, creating problems that are trivial once effectively divided into parts. Performance remains poor in participants with low fluid intelligence, but is radically improved by problem layout that aids cognitive segmentation. In line with the principle of compositionality, we suggest that effective cognitive segmentation is important in all organized behavior, explaining the broad role of fluid intelligence in successful cognition.

fluid intelligence | problem solving | cognitive compositionally | focused attention

It is widely argued that the power of human cognition rests heavily on the principle of compositionality, or the ability to build indefinitely complex mental structures from the organized combination of simple parts (e.g., refs. 1–3). In this article, we link this idea of compositionality to the psychometric concept of fluid intelligence. In psychometrics, fluid intelligence is conventionally measured with tests of novel problem-solving, such as Raven’s Progressive Matrices (4) or Cattell’s Culture Fair (5). Such tests derive their importance from broad correlations with cognitive success across many different kinds of tasks and settings. Here we argue that the core ingredient is closely related to the cognitive principle of compositionality.

There have been several influential proposals concerning core cognitive factors in fluid intelligence. One popular hypothesis suggests that fluid intelligence reflects the capacity of working memory (6), whereas in a second hypothesis, fluid intelligence reflects the speed of processing (7). Indeed, fluid intelligence tests show some positive correlation with working memory or speed tasks, as they do with almost any task in a cognitive battery (8). That said, it is complex, multipart tests that show the broadest pattern of strong positive correlations across many different tasks (8–10). In matrix tests, for example, the results of multiple cognitive steps must usually be combined to determine each item solution. In tests of this sort, it seems likely that complexity itself is critical (11, 12).

Consistent with this argument, several findings link fluid intelligence to “executive control” functions of the frontal lobe (e.g., refs. 8, 13–15), or a more distributed network comprising regions of lateral frontal, insular, dorsomedial frontal, and parietal cortex (11, 16). Performance of traditional fluid intelligence tests is associated with extensive activity within this network (17, 18), and sensitive to lesions affecting it (ref. 19; see also ref. 20). Recently, we have linked the function of this frontoparietal control network to the broad principle of cognitive compositionality. Early work in artificial intelligence established the importance of dividing complex problems into simpler, more manageable parts (e.g., ref. 21). A high-level goal, for example, is generally achieved by hierarchical division into a complex structure of subgoals, with successive focus on each part of the problem in turn (e.g., ref. 22; see also, ref. 23). If this is not done, behavior can become unstructured and chaotic (24), resembling the chaotic behavior typical of frontal lobe patients (25), especially in complex, unstructured situations (see, e.g., refs. 26 and 27). Following this work, we have proposed that the core function of the distributed frontoparietal executive control system is one of cognitive segmentation, or dividing complex behavior into a series of separate, simpler parts (11, 28). Such segmentation implies using knowledge of a task domain to focus attention on useful task parts, producing a structured mental control program. Cognitive segmentation, we suggest, is required in any organized behavior, but is especially important in novel, multistep tasks such as Progressive Matrices, in which a new structure of attentional episodes must be discovered and created for each new problem.

With its emphasis on focused attention, our proposal has similarities to others that link low fluid intelligence to less focused or targeted cognition (14, 15, 29). Consistent with a core role of frontal cortex in creating attentional episodes, in the behaving monkey, lateral prefrontal cortex shows dynamic neural activity as a task progresses, with selective emphasis of information relevant to a current cognitive step (e.g., refs. 30–32) and radical reorganization of activity from one task step to the next (33, 34). In line with similar patterns of frontal and parietal activity shared by many different tasks (35, 36), these results suggest a highly adaptive neural medium, constantly reorganizing to foreground information relevant to current thought or behavior (11, 37).

In this article, we contrast a segmentation account of fluid intelligence with accounts focusing on working memory capacity and mental speed. To this end, we modify traditional matrix problems, aiming to make segmentation easy or difficult to achieve and, at the same time, eliminating any major role for other factors.

Significance

Tests of fluid intelligence are important for their broad association with effective cognition and lifetime achievement. An enduring question concerns basic cognitive mechanisms measured in such tests. Fluid intelligence is usually measured with complex problem-solving tasks, and in such tests, we suggest that the core limit is one of cognitive segmentation, or managing complex activities by selective attention to separate, simpler parts. Here we modify traditional fluid intelligence problems to test this hypothesis and to minimize the roles of working memory capacity and mental speed. The findings suggest a cognitive interpretation for what it is that fluid intelligence tests measure, based on dynamic attentional control functions of frontal and parietal cortex.

Author contributions: J.D., D.C., D.J.M., and A.B. designed research; J.D., D.C., and D.J.M. performed research; J.D., D.C., and D.J.M. analyzed data; and J.D. and D.J.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

To whom correspondence should be addressed. Email: john.duncan@mrc-cbu.cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1621147114/-/DCSupplemental.
An example of a matrix problem in typical format is shown in Fig. 1. In this problem, the task is to decide which of the four response alternatives at the bottom completes the matrix at the top. To determine the correct solution, it is necessary to take account of three varying stimulus features: whether the top part is outline or black, whether the left part is curved or angled, and whether the right part is straight or bowed. Only by considering all three features can the correct solution be determined, and reflecting the importance of complexity, if the problem has fewer varying features, it becomes progressively easier to solve (10).

In a problem like this, increasing the number of varying features has several consequences. First, each component feature must be identified and the correct value determined. Second, solutions to one part of the problem must be held in working memory whilst working on others. Third, the different parts of the solution must be integrated to allow a final choice among response alternatives. In their classic analysis of this problem, Carpenter et al. (10) give a central role to maintaining and integrating complex information in working memory. Integration is also important in some accounts of task complexity based on speed: Speed may be important in complex tasks because combining the results of different task operations requires that they all be available at the same time (7). Here, we wished to minimize load on integration, working memory, and speed. We modified the task such that the only significant requirement was to break the three-feature problem into appropriate one-feature parts; that is, to focus attention on one soluble part after another. We predicted, nevertheless, substantial difficulties in participants with low fluid intelligence, and that these difficulties would largely be removed by cues making appropriate segmentation easy to achieve. To further assess the role of speed, we used two task versions: one giving limited time for each object part, and the other with no time limit.

**Results**

Our task modifications are illustrated in Fig. 2. As in the traditional problem in Fig. 1, each matrix used objects with three varying parts (Fig. 2A, combined format). Now, however, the participant was provided just with a single answer box, and was asked to draw the correct answer within this. There was thus no requirement to store one part of the answer in working memory while working on others, or to integrate the three parts into an overall mental image of the correct solution. Instead, participants could focus on each part of the matrix objects in turn, work out the correct solution for this part, and draw it immediately into the response box. Note that, in any novel problem, there is always ambiguity over how materials should be represented or described to break them into useful parts. By constructing matrix entries from 3 clearly distinct parts, we attempted to make the appropriate segmentation as transparent as possible (e.g., Fig. 2, separation into left-side shape, right-side arrow, and vertical line; for full set of materials, see Fig. S1). Despite these changes to traditional task format, we predicted that substantial difficulties would remain for participants with low fluid intelligence. To confirm that cognitive segmentation was the critical difficulty, we also introduced a condition in which this process was made trivially easy. In this condition (separated format, Fig. 2B), separate matrices were presented for each object part. The task was otherwise unchanged; again, the participant had to focus on each part in turn and draw the solution for this part into the single response box, with a complete three-part answer finally built up as in the combined-format condition. Now, however, we predicted that errors would largely be eliminated.

To test these predictions, each participant was given the two conditions of our new matrix task, along with the Culture Fair measure of fluid intelligence (5). Each condition of the matrix task began with two practice trials, leading the participant through the process of focusing on and drawing one object part at a time. Practice was followed by 10 scored trials (Fig. S1). In Experiment 1, 40 participants received tests in traditional paper-and-pencil format, with a time limit of 30 s per problem. In Experiment 2, 21 new participants drew their answers instead on an electronic tablet, allowing detailed measurement of response timing. To further assess the importance of speed, participants in Experiment 2 were given unlimited time to complete each problem.

For Experiment 1, scatterplots relating proportion of correct answers in the matrix task to Culture Fair IQ are shown in Fig. 3A. For the combined-format problems, participants with low Culture Fair IQ showed very poor performance. With the separated format,
in contrast, most items were solved correctly across the fluid intelligence range. Despite unlimited time to solve each problem, and some resulting improvement in performance, the result was replicated in Experiment 2 (Fig. 2B). The data were examined using the general linear model, predicting proportion of correct answers from Condition (combined or separated), Experiment, and IQ. The main effects of Condition \(F(1, 57) = 26.8; P < 0.001\) and IQ \(F(1, 57) = 14.9; P < 0.001\) were both highly significant, along with their interaction \(F(1, 57) = 12.0; P = 0.001\). Despite the trend for improved performance in Experiment 2, Experiment showed no significant main effect, \(F(1, 57) = 3.1; P = 0.08\) or interactions.

Combining data across experiments, proportion correct in the combined-format condition showed a partial correlation (Pearson’s \(r\), with effect of Experiment partialled out) of 0.52 with Culture Fair IQ, in line with very poor performance for the low-IQ participants. For the separated format, the few errors remaining also tended to be made by low-IQ participants \((r = 0.33)\).

As the Culture Fair has 4 subtests (series, odd-one-out, matrices, topology), we were able to examine any possible influence of problem type. For the combined condition of our modified matrix task, partial correlations with Culture Fair subtests (removing the effect of Experiment) were 0.45 (series), 0.38 (odd-one-out), 0.35 (matrices), and 0.40 (topology), suggesting a broad link to fluid intelligence, rather than specific overlap with the Culture Fair’s own matrix problems. We also compared our integrated matrices to the Culture Fair’s own matrices in terms of correlation to remaining Culture Fair subtests (sum of series, odd-one-out, and topology.) Intriguingly, the partial correlation with remaining subtests was somewhat higher (0.53) for our modified problems than for the Culture Fair’s own matrices (0.41).

Although practice trials already illustrated the procedure of focusing on one object part after another, we examined whether problem-solving in the integrated condition would be helped by prior experience of the separated condition, perhaps further reinforcing part-by-part attentional focus. Performance in the integrated condition, however, was independent of whether it was experienced first or second \(F(1, 53) = 0.2\).

Additional insight into problem-solving failures was provided by a detailed analysis of drawing errors. In Experiment 1, for combined-format problems, pooling across participants and items, a total of 289 parts were not correctly drawn. In 149 cases (52%), the participant drew the wrong one of the two alternative values given in the matrix (wrong-alternative errors). In addition, 83 cases (29%) were omissions of a part, with a variety of other incorrect drawings making up the remaining 57 cases. For separated-format problems, a total of 50 parts were not correctly drawn, with 42% wrong-alternative errors, 32% omissions, and the remainder miscellaneous. In Experiment 2, for combined-format problems, there were 67 wrong-alternative errors and 7 omissions (79% and 8%, respectively) among the total of 85 cases in which a part was not correctly drawn. For separated-format problems, the total of 12 errors was made up of 8 wrong-alternative errors and 4 omissions. Although some errors in Experiment 1 likely reflected failure to complete the problem in the time available, the majority throughout were confusions between correct and incorrect solutions for a given object part.

In Experiment 2, we had access to drawing times for each stroke of the participant’s solution. These data allowed us to confirm that, as expected, participants predominantly focused on one object part a time, with long pauses between drawing one part and the next. Time from problem presentation to first stroke was substantially longer for the combined-format condition (mean \(= 13.3\) s) than for the separated-feature condition [mean \(= 7.3\) s; \(t(15) = 4.8; P < 0.001\); data unavailable for 5 participants because of a procedural error]. Total time spent drawing (time from first to last stroke), in contrast, was similar in the two conditions [22.7 and 22.6 s, respectively; \(t(20) < 0.1\)].

Excluding the few cases in which a single object part was not drawn as a whole before starting the next (10.1% and 2.8%, respectively, for combined- and separated-format problems), mean times to draw a single object part were 3.1 and 2.6 s, respectively, for combined- and separated-format \(t(20) = 1.6; P > 0.05\), with mean pauses between the end of one part and the start of the next of 7.1 and 7.6 s, respectively \(t(20) = 0.8; P > 0.05\). The data show closely similar solution strategies in the two conditions, with each part of the solution drawn before moving on to consider the next.

**Discussion**

Matrix problems are among the most widely used tests of “fluid intelligence.” They are important because ability to solve these problems is broadly predictive of success in many kinds of cognitive activity. The critical cognitive ingredient of such problems remains uncertain. To address this question, we made a number of simple modifications to the traditional matrix format. Straightforward though they are, these modifications put major constraints on understanding what a matrix test measures.

In particular, we aimed to link fluid intelligence to the broad principle of cognitive compositionality and to the attentional control functions of frontal and parietal cortex. The key process, we propose, is one of splitting a complex whole into simple, separately attended parts. To contrast with influential views based on working memory or mental speed, we modified the matrix format to minimize working memory and speed demands. By constructing matrix items from multiple parts and allowing answers for each part to be drawn in turn, we removed the requirement to store intermediate results and finally synthesize into a single answer. We
also used both speeded and unspeeded task versions. Despite these modifications, performance remained very poor in participants with low fluid intelligence. Among the many errors made, the most common was choice of the wrong alternative value for a given part, implying confusion in solving this aspect of the problem. Such errors largely vanished, however, when the materials made it trivial to separate the overall problem into parts. Of course, such data do not show that working memory capacity and/or speed make no significant contribution to fluid intelligence. Even when little remains in a matrix problem beyond the need to split it into easily solved parts, it appears still to capture the essence of traditional tests.

As addressed in the long history of symbolic artificial intelligence (e.g., ref. 22), splitting a problem into parts must be based on knowledge of the task domain, in the present case including knowledge of objects, matrices, tasks, rules, and so on. Attentional focus must be achieved by using this knowledge to discover important parts of a problem, or component steps that move closer to the overall goal. In the present tasks, this would correspond to focus on useful component parts of the objects depicted in the matrix. Plausibly, knowledge is widely distributed in the brain, with frontoparietal control systems important in selecting and combining together the perceptual, memory, and action components of a current attentional episode (38).

Even the simplest tasks generally have some correlation with fluid intelligence, and in the current experiments, even performance in the separated condition correlated with the Culture Fair. This is the result we should expect, as even in simple tasks, attention must be focused on right things at the right time, producing an appropriate mental control program. In a typical laboratory task, for example, components might include ensuring appropriate fixation and readiness before a stimulus is presented, performing whatever operations on that stimulus the task requires, monitoring response timing and accuracy, and so on. This universal requirement for building a complex whole from focused parts may be at least one major explanation for the finding of universal positive correlations between fluid intelligence and even simple tasks. As tasks become more complex, however, it is increasingly challenging to separate them into clearly focused parts. The best way to measure cognitive segmentation may be with complex, multistep behavior, such as the problem-solving of traditional fluid intelligence tests.

Cognitive segmentation implies focused attention on separate parts of a complex problem, and many observations support the central role of this process in effective thought and behavior. Classical accounts of frontal lobe damage, for example, emphasize disorganization in sequences of behavior, without a series of steps clearly leading to the goal (25). In plans for everyday activities, such as instructions for self-assembly furniture, much use is made of bullet points and similar devices to create a useful division into parts. In adults’ interaction with young children, “scaffolding” of effective behavior is useful only when it divides complex tasks into simpler, manageable parts (39). More generally, “abstraction,” long held to be a critical aspect of frontal lobe function (40), by definition involves focused attention just on some selected aspect of a complex whole, usually the aspect that is useful for some cognitive purpose. Cognition in general is organized in a structure of focused parts; as Lashley (41) foreshadowed, understanding such structure may be an essential step toward a “physiology of logic” (41, p. 122).

Materials and Methods

Experiment 1. Participants. Forty participants (mean age, 57.3 years; range, 41–71 years; 25 female) were recruited from the volunteer panel of the MRC Cognition and Brain Sciences Unit. Participants gave informed, written consent and were reimbursed for their time. All procedures were carried out in accordance with ethical approval obtained from the Cambridge Psychology Research Ethics Committee.
Access to Data and Materials. Materials, code, and data are freely available from the authors on request. The full set of matrix problems is provided in Fig. 51.

ACKNOWLEDGMENTS. This research was supported by Medical Research Council intramural program MC-A060-5PQ10.

5. Institute for Personality and Ability Testing (1973) Measuring Intelligence with the Culture Fair Tests (The Institute for Personality and Ability Testing, Champaign, Illinoi).