

Remembering What We Learn

By Henry L. Roediger, III, Ph.D., and Kathleen B. McDermott, Ph.D.



Source/Shutterstock

Editor's Note: Memories are the internal mental records that we maintain, which give us instant access to our personal past, complete with all of the facts that we know and the skills that we have cultivated. While the mind's capacity to store and recall information is truly wondrous, there are desirable and undesirable difficulties in learning. Our authors provide examples of retrieval practice and individual differences in long-term retention and explore quick and slow learners.

Children are born learners, picking up information about the world shortly after birth and continuing for the rest of their lives. We might expect them to be experts when they arrive at school, eager to learn and finding it easy. Yet every teacher knows that things are not so simple. Students flounder with many ideas and concepts, and no wonder. They are being asked to acquire knowledge that, in some cases, took centuries of careful thinking and research to discover, and that is often counterintuitive.

In the past 20 years, cognitive psychologists have produced dozens of experiments in which students learn material two different ways and are asked to choose which one they learn better with. Based on their performance results, they often choose wrong, predicting a difference when the variable has no effect, predicting no effect when the variable has one, or even mistaking the direction of the effect—they predict one condition will lead to better learning than a second, yet the second condition produces superior performance.

An example is having college students learn a list of words presented on a computer, with half the words in a large font and the other half in a small font. Students predict they will recall the words in the large font better than those in the small font, but actually the variable produces no detectable difference; students recall them equally well.²

Our article is about the more perplexing case where students predict that one variable will enhance long-term retention, yet the other leads to better performance. Three variables that have this effect are spacing versus massing practice sessions on a task; interleaving or mixing up the type of skills one performs (e.g., practicing different kinds of math problems in a set) rather than blocking them (repeated practice on one type of problem, then the next, and so on); and retrieval practice—recalling information while learning—relative to rereading it (what most students report doing). Students predict that massing practice sessions, using the blocking method, , and rereading will lead to better long-term retention than the spacing, interleaving method, or retrieval practice, yet in each case the opposite occurs.¹ As we will discuss, one reason that students (and teachers) have trouble discovering strategies that lead

to durable long-term learning is that sometimes the strategies that enhance short-term performance lead to poor long-term retention. Cramming for a test (repeated studying just before the test) works fine for an immediate test, but information learned this way is rapidly forgotten.³

Desirable Difficulties in Learning

In 1994, Robert Bjork coined the term *desirable difficulties in learning* to capture a set of paradoxical findings that continue to fascinate researchers today.⁴ The concept, though descriptive, helps us to understand the wrong intuitions of students about the effectiveness of learning strategies. The basic finding, foreshadowed in the previous section, comes from experiments that vary a method of learning and then report a strong (cross-over) interaction between performance on an immediate test and on a delayed test. That is, Condition A produces better performance than Condition B when the test is given soon after learning, but if the test is delayed even a day or two, Condition B leads to superior performance. Thus, the condition that at first seems to work more poorly produces better learning for the long term: the difficulty that condition B initially seems to present is advantageous in the long run. In education, of course, performance is often assessed soon after learning, and so strategies that enhance short-term performance are generally assumed to be good. But they aren't, or at least not always.

The three variables briefly mentioned above are known to produce this pattern of results. Spacing of practice is one: Its benefit was discovered by the German psychologist Hermann Ebbinghaus in 1885 and has been replicated countless times. If people are given a passage to learn and are asked to read it twice, massed practice (back-to-back readings) will lead to better recall on an immediate test than will spaced practice (reading the passage twice with time between devoted to other activities). Cramming, in other words, works in the short term. But if the test is delayed for an extended period, the opposite pattern occurs: spaced reading leads to better recall.⁵

A related issue to spacing is interleaving, and it has a similar effect to spacing. For example, if students are learning to solve geometry problems, the typical method of instruction has them learn to do one type of problem (e.g., finding the volume of a particular type of solid such as a

wedge) with repeated practice, then addressing the same problem with another shape, such as a spherical cone. And so on. Thus, instruction on finding volume of solids is blocked by the type of solid in classroom learning and in homework in math classes. In fact, this type of instruction is almost universal in math instruction—learn and master one type of problem at a time.

A different way to teach from the blocked method just described is called interleaving of problems (also known as “shuffling” of problems in math education). Here, students learn the rules for all types of solids and then intersperse their practice with different examples, so they solve for one type of problem, then a different one, and so forth. Not surprisingly, students find this method hard, and learn more slowly. On a test given immediately after instruction, students who learn by interleaving solve new problems more poorly than those trained with blocked practice. However, when students are tested a week later, the ones who learned by interleaving show little forgetting, whereas the performance of those who learned by blocked methods declines precipitously.⁶ Thus, at a delay, interleaved practice leads to better performance than blocked practice.⁷ The reason is that with interleaved practice, students learn to discriminate among types of problems and to select the right method for each, while in blocked practice they know the type of problem at the outset, and so get no practice at learning to choose which approach is appropriate for the case at hand⁸.

A third variable exhibiting desirable difficulty is retrieval practice, i.e. recalling or recognizing information when tested (either by testing oneself or by a teacher in a classroom). In one experiment (there have been hundreds), Roediger and Karpicke³ compared repeated testing of recently studied information to repeated studying of the same material (what students usually report that they do).⁹ They gave college students a short passage to read about the sun and sea otters. In one condition, students followed their initial study of the passage with three additional study periods (SSSS, where S is for study). In a second condition, students studied the passage two additional times and took one test (SSST). A third group studied the passage only once and then took three successive tests (STTT).

The tests were free recall, which means that students were told to recall everything they could from the passage they had read. The subjects who were tested during the learning period recalled about 70 percent of the units of information on each of the tests, but students who

reread the passage were, of course, re-exposed to 100 percent of the ideas on each reading. After these manipulations, students were asked to estimate how well they would remember the passage a week later. To assess their actual retention, they were given a final free recall test, with half the students taking the test only five minutes after the last trial in the study schedule and the other half taking it a week later.

As shown below in Figure 1, when tested five minutes after learning, students who had studied the passage four times (SSSS) recalled about 12 percent more content than those who had studied the passage once followed by three tests during the learning period (STTT); that is, there was an advantage to repeated studying. However, when the final assessment test was given after a weeklong delay, the pattern reversed: repeated testing (STTT) during the initial learning period produced 21 percent greater recall than repeated studying (SSSS). Recall was intermediate in both the immediate and delayed test among students in the single test condition (SSST).

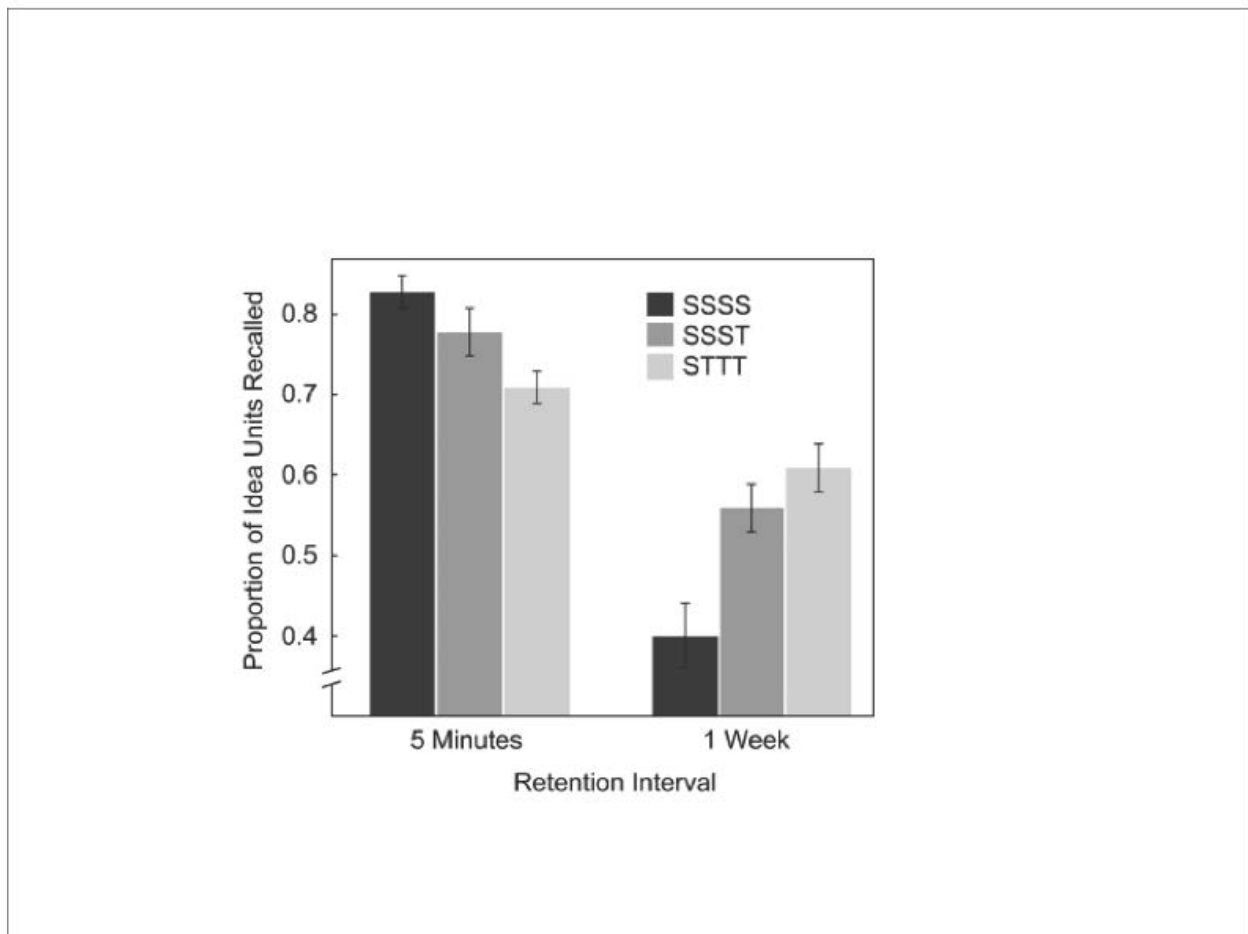


Figure 1. Mean proportion of idea units recalled following repeated studying or repeated testing during a learning phase. The final criterial test occurred either 5 minutes or one week after learning. Repeated (massed) studying led to best performance on the immediate test, but increased testing caused better recall on the delayed test. Figure 2 from Roediger and Karpicke.³

This effect can be accounted for by relative forgetting between the immediate and delayed tests: after a week, participants forgot nearly half of what they could initially recall when they had studied the text four times (SSSS), but only forgot about 14 percent of the content when they had studied it once and took three tests during learning (STTT). Importantly, the initial tests in this experiment were not followed by feedback about what students missed; if such feedback had been provided, the effect would have been even larger. The primary take-home message from this experiment is that retrieval practice, here in the form of tests, enhances long-term retention of information by slowing the rate of forgetting.

The enhanced retention following repeated retrieval practice is especially impressive given that the students spent considerably more time exposed to the ideas in the repeated study schedule than with repeated testing, so they had more opportunities to learn the material. Furthermore, students were unaware of the benefits of testing. In the judgment they made just after the learning conditions, students who had repeatedly read predicted they would recall more a week later relative to those who had studied the passage only once and been tested three times. In short, their predictions were exactly the opposite of the actual recall pattern. Why were they so wrong?

Perhaps learners are not skilled at predicting what they will forget, but instead base their predictions on what they know when they make them. After all, their predictions were accurate regarding the tests given soon after learning. In other words, they took current performance to be an indicator of long-term learning, and this mistake creates the illusion of mastery that is a hallmark of poor training techniques. The same process underlies erroneous prediction seen in spacing and interleaving studies.

The benefits of retrieval practice are well documented,^{10,11} but its neural substrates are not yet well understood. Although a small literature has begun to emerge in this area, different labs

have studied the problems in various ways, so conclusions are difficult; an early review of this work¹² concluded that there was little convergence in the findings.

Individual Differences as a Possible Desirable Difficulty

A fascinating puzzle is why some people learn faster than others and how this affects retention of the information learnt. Aristotle likened memory to a block of wax onto which experience made impressions, and he contended that neither unusually fast nor unusually slow learners were likely to remember well; rather, the best would be those who learned at an intermediate rate. This, he conjectured, is due to the quality of their “wax,” neither too soft nor too hard to make a good impression. But for the extremes, in his words, “the design makes no impression ... because of the hardness of that which is to receive the impression ... For a similar reason, neither the very quick or the very slow appear to have good memories; the former are moister than they should be, and the latter harder; with the former the picture has no permanence, with the latter it makes no impression.”¹³

A basic question for researchers in human learning and memory but one generally overlooked is whether fast, intermediate, and slow learners really do differ in this way. If people who learn at different speeds are all brought to point where they have learned the information perfectly, will their loss of information over time (i.e. forgetting) drop off similarly or (as Aristotle would have it) to different degrees? Might slow learning actually be a desirable difficulty? Might people who take longer to get information into memory be able to hold on to it well because they have had much more practice?

Subjects in one study were given 45 foreign-language vocabulary word pairs to learn (e.g., NAMAS-HOUSE).¹⁴ After studying each pair once, people took successive tests with feedback (the correct answer was supplied whether the subject’s answer had been right or wrong). Each successive test consisted only of items that had not previously been retrieved. Speed of learning was indexed by the number of tests a person had to take before all items had been correctly recalled. Hence, faster learners took fewer tests. Would they also forget faster than ones who took more tests (the ones in the middle of the pack)? The answer is *no*.

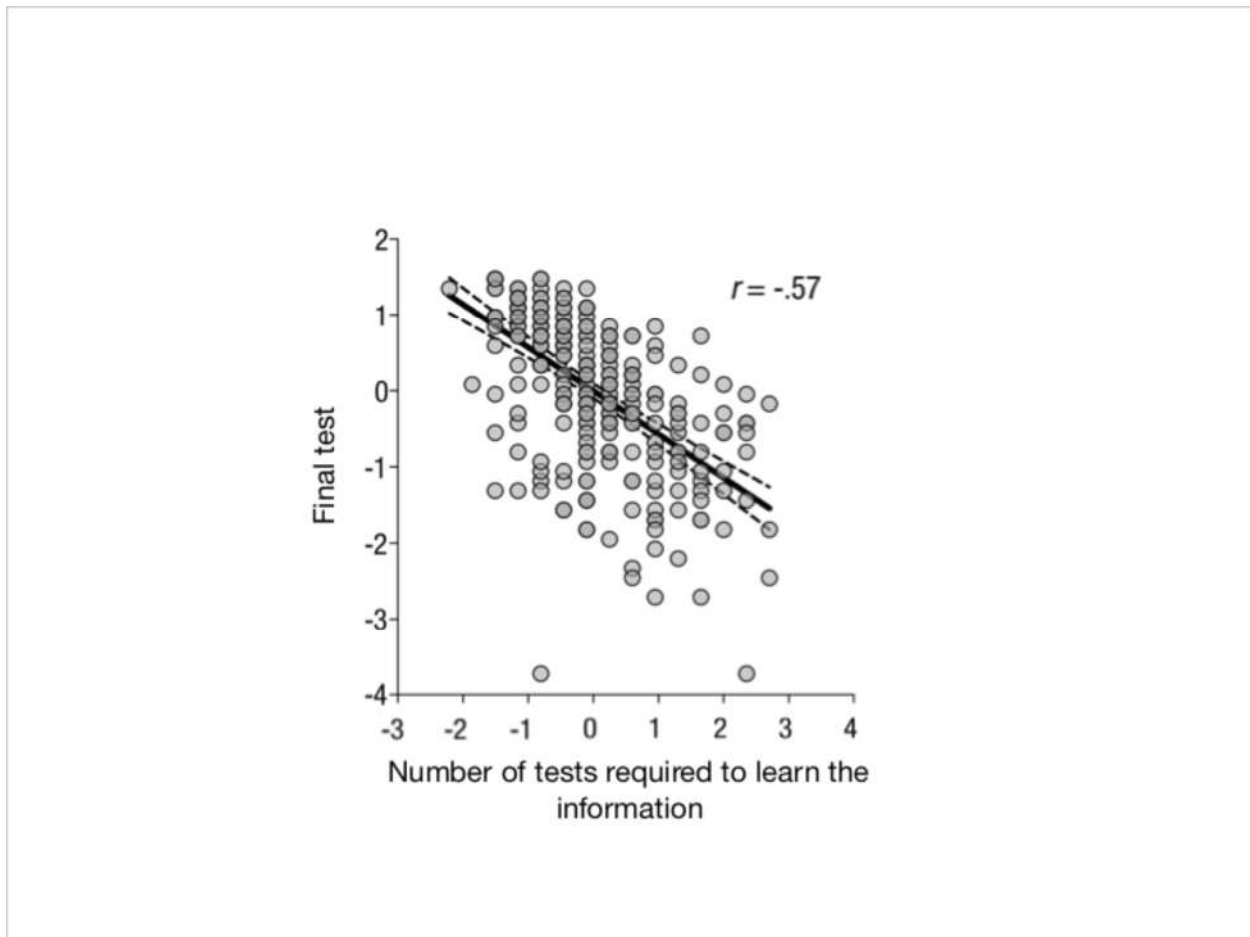


Figure 2. Scatterplot (with best-fitting regression line and 95 percent confidence interval) showing a consistent relation between learning speed and retention across people. Subjects who learned 45-word pairs more quickly (fewer Tests to Criterion) tended to recall more word pairs on a delayed cued recall test (higher Final Test Scores). Data are plotted in standardized scores. Figure 4 from Zerr et al.¹⁴

Figure 2 shows the relation between final recall (on the ordinate) as a function of the number of test trials to reach the criterion of 100 percent correct (represented on the abscissa), with the data reported in standardized scores (such that 0 is average and higher scores are better scores). Keep in mind that low numbers of test trials indicate faster learning. As can be seen in Figure 2, people who learned more quickly (i.e., took fewer tests to learn the material) tended to also retain more items after a delay. That is, the difference is in learning efficiency: fast learners are slow forgetters (and slow learners are fast forgetters). From this analysis, we can conclude that being a slow learner is not a desirable difficulty. Even though slower learners get much more practice when learning the material, they forget it faster.

How can we reconcile this finding with the theory of desirable difficulties? Is there a contradiction? On the surface, yes: being a slow learner is not a desirable difficulty. However, the desirable difficulties framework was not designed to capture differences between people; rather, it was intended to explain why, for the same individual, more effortful learning conditions can be beneficial. And the word *desirable* is in the phrase for a reason; not all difficulties are desirable ones.⁴ Being a slow learner is not desirable.

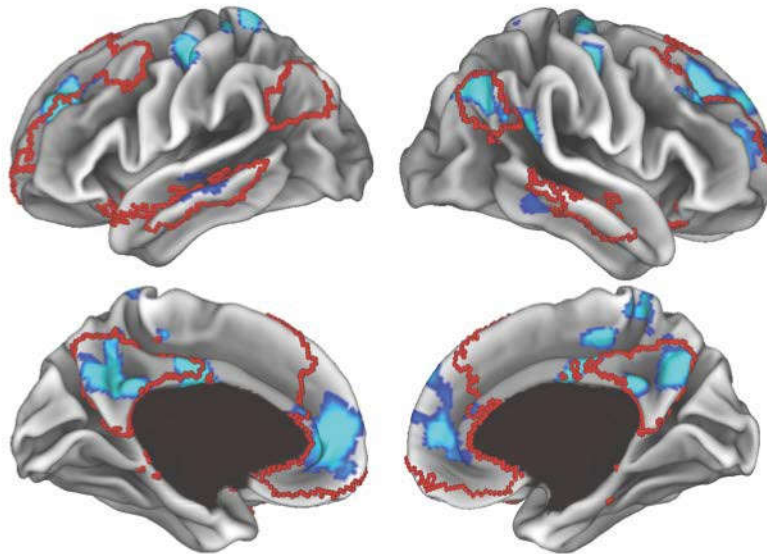


Figure 3. Across 86 participants, those who are more efficient learners tend to have greater deactivation in the regions shown in blue than less efficient learners¹⁵. These regions correspond to the default mode network, as outlined in red (as defined in ¹⁸). Adapted from Figure 4 from Nelson et al.¹⁵

We can look to neuroimaging evidence to help understand these differences between fast and slow learners. During the initial study period, people with higher learning efficiency show a characteristic pattern of results within ventromedial prefrontal cortex, posterior cingulate cortex, superior frontal gyrus, and angular gyrus, as can be seen in Figure 3.¹⁵ These regions collectively form a brain network known as the default mode network.¹⁶ The network's name derives from the observation that these regions tend to be active during our typical (or default)

modes of thinking when we are not performing a task (that is, when we are daydreaming, or letting our minds wander). This network becomes inactive when we focus our attention away from our default, self-focused thoughts and instead focus on the external world.¹⁷ In the learning efficiency study, people who were efficient learners tended to deactivate this brain network more strongly during original encoding than did slow learners, who either deactivated weakly or not at all.¹⁵ One way of thinking about this finding is that highly efficient learners can more effectively set aside the natural proclivity for self-focused thought (e.g., daydreaming) and turn their attention to the task at hand.

Students (and teachers) often believe that strategies that produce good student performance in the short term are generally excellent learning techniques. That attitude is perfectly understandable; a good example is cramming, which may get the students through an immediate test with good grades. However, the results we have presented show that often strategies that seem to slow or hamper learning in the short term actually aid recall in the long term. Three variables—spaced practice, interleaved practice, and retrieval practice—impair learning on an immediate test but lead to better performance on a test after a longer delay.

Because students and teachers often focus on short-term effects, which they can easily observe, they may fail to recognize the good strategies that can be used for long-term retention. Interestingly, learning speed or learning efficiency represents as an exception. Being a less efficient learner is an undesirable difficulty. Slow learners, despite having much greater amounts of practice in learning to reach a 100 percent level of recall, forget the learned information more rapidly. Nonetheless, if those slower learners employ the good strategies like retrieval practice for learning, they can greatly improve their learning and retention.

References

[1] M.G. Rhodes, A.D. Castel, Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, 137 (2009), 615-625.

- [2] R.A. Bjork, J. Dunlosky, N. Kornell, Self-regulated learning: Beliefs, techniques, and illusions, *Annual Review of Psychology*, 64 (2013) 417-444.
- [3] H.L. Roediger, J.D. Karpicke, Test-Enhanced Learning: Taking memory tests improves long-term retention, *Psychological Science*, 17 (2006) 249-255.
- [4] R.A. Bjork, Memory and metamemory considerations in the training of human beings. , in: J. Metcalfe, A.P. Shimamura (Eds.) *Metacognition: Knowing about knowing* MIT Press, Cambridge, MA, 1994, pp. 185-205.
- [5] S.K. Carpenter, Spacing effects on learning and memory, in: J.T. Wixted (Ed.) *Cognitive Psychology of Memory*, Academic Press, Oxford, 2017.
- [6] D. Rohrer, K. Taylor, The shuffling of mathematics practice problems improves learning, *Instructional Science*, 35 (2007) 481-498.
- [7] S.H.K. Kang, The benefits of interleaved practice for learning, in: J.C. Horvath, J. Lodge, J.A.C. Hattie (Eds.) *From the laboratory to the classroom: Translating science of learning for teachers*, Routledge, NY, 2017.
- [8] D. Rohrer, Interleaving helps students distinguish among similar concepts, *Educational Psychology Review*, 24 (2012) 355-367.
- [9] J.D. Karpicke, A.C. Butler, H.L. Roediger, Metacognitive strategies in student learning: Do students practise retrieval when they study on their own?, *Memory*, 17 (2009) 471-479.
- [10] C.A. Rowland, The effect of testing versus restudy on retention: A meta-analytic review of the testing effect, *Psychological Bulletin*, 140 (2014) 1432-1463.
- [11] O.O. Adesope, D.A. Trevisan, N. Sundararajan, Rethinking the use of tests: A meta-analysis of practice testing, *Review of Educational Research*, advance online publication, DOI 0034654316689306(2017).
- [12] P. van den Broek, A. Takashima, C. Wiklund-Hornqvist, L.K. Wirebring, E. Segers, L. Verhoeven, L. Nyberg, Neurocognitive mechanisms of the “testing effect”: A review, *Trends in Neuroscience and Education*, 5 (2016) 52-66.
- [13] F. Yates, *The art of memory*, The University of Chicago Press, Chicago, 1966.
- [14] C.L. Zerr, J.J. Berg, S.M. Nelson, A.K. Fishell, N.K. Savalia, K.B. McDermott, Learning efficiency: Identifying individual differences in learning rate and retention in healthy adults, *Psychological Science*, in press (2018).

- [15] S.M. Nelson, N.K. Savalia, A.K. Fishell, A.W. Gilmore, F. Zou, D.A. Balota, K.B. McDermott, Default mode network activity predicts early memory decline in healthy young adults aged 18-31, *Cerebral Cortex*, 26 (2016) 3379-3389.
- [16] M.E. Raichle, A.M. MacLeod, A.Z. Snyder, W.J. Powers, D.A. Gusnard, G.L. Shulman, A default mode of brain function, *Proceedings of the National Academy of Science USA*, 98 (2001) 676-682.
- [17] G.L. Shulman, M. Corbetta, J.A. Fiez, R.L. Buckner, F.M. Miezin, M.E. Raichle, S.E. Petersen, Searching for activations that generalize over tasks, *Human Brain Mapping*, 5 (1997) 317-322.
- [18] J.D. Power, A.L. Cohen, S.M. Nelson, G.S. Wig, K.A. Barnes, J.A. Church, A.C. Vogel, T.O. Laumann, F.M. Miezin, B.L. Schlaggar, S.E. Petersen, Functional network organization of the human brain, *Neuron*, 72 (2011) 665-678.