Improving intelligence: a literature review

Martin Buschkuehl, Susanne M. Jaeggi
University of Michigan, Ann Arbor, USA

Summary

Intelligence is associated with performance on a wide range of cognitive tasks and is a reliable predictor for educational and professional success. Therefore, the development of effective training regimens that aim to improve intelligence is of high interest. In recent years, there has been accumulating evidence that certain interventions have a positive impact on intelligence. The aim of the present paper is to provide a literature review on such studies. Despite promising results, we conclude that there are still many unknown variables and that the mechanisms that underlie improvements in intelligence are not well understood. More work is needed in order to disentangle these involved processes and to further refine existing training regimens.

Key words: intelligence; training; transfer; working memory; executive functions

Introduction

In his seminal paper on intelligence published more than hundred years ago, Spearman [1] found that people who performed well in one intellectual domain also performed well in others. He concluded that this positive correlation favored the existence of a general factor $G$ that is common to all tests of intellectual ability. Today, the term ‘intelligence’ is used variably [2]; however, most researchers seem to agree on the ability to learn being a central aspect of intelligence. This is in line with findings that $G$ is a very good predictor of academic achievement [3]. Prominent theories of intelligence divide $G$ into two components: crystallized intelligence ($Gc$) and fluid intelligence ($Gf$) [4]. $Gc$ refers to knowledge acquired by past experience, for example, vocabulary or skills. $Gc$ can be measured with vocabulary tests or tasks requiring general knowledge. In contrast, $Gf$ is the ability to cope with new situations for which previously acquired knowledge is only minimally helpful [e.g., 5]. Also, $Gf$ has been commonly regarded as the most reliable and predictive measure for successful performance in both educational and professional settings [3, 6–8]. As there is also a lot of empirical evidence showing $Gf$ as the best predictor for a wide variety of tasks [9], $Gf$ is conceptually very close to $G$ [4, 10]. Prototypical tasks to measure $Gf$ are so-called matrix reasoning tasks [11], such as Raven’s Progressive Matrices [12], which is one of the most frequently used tests. In such matrix reasoning tasks, the participant is presented with a pattern of logically related pieces. One piece of the pattern is missing and the participant is instructed to pick the piece that logically fits into the empty slot by selecting the correct one out of several possibilities. It is very easy to increase performance in such tests by simply practicing the tests themselves [13]. This is even the case in old adults where neural plasticity is assumed to be reduced [14, 15]. Nevertheless, it has been demonstrated that practice on these tests decreases their novelty and with that the underlying $G$-related processes, and furthermore, the predictive value of the tests for other tasks is largely reduced [16, 17]. The question of interest is whether it is possible to increase $G$ or $Gf$ per se, that is, not by practicing intelligence tests themselves, but by improving fundamental processes or prerequisites that form the basis of intelligent behavior. This question is by no means a new one and was discussed by Jensen [18, 19] but also by others more recently [20]. Until a couple of years ago the general conclusion of these discussions was that interventions aiming to improve intelligence resulted in only very little if any success at all [21].

But in recent years, a growing number of studies has been published showing that certain interventions have indeed a positive impact on some measures of intelligence. The reason for this recent development is most likely due to the advances in cognitive-based theories of intelligence that provide insights into what kind of training might be successful to promote intelligence [22]. The aim of this contribution is to provide a brief overview of studies that showed improvement of intelligence after some form of intervention. The reviewed studies are divided into two groups: studies that used intervention approaches that are focused on training of working memory (WM) and executive functions, and secondly, studies which entailed other approaches.
Improving intelligence by training on working memory and executive functions

The overall rationale behind the first group of studies is based on the observation that WM is closely related to intelligence measures [23–27]. As it is assumed that WM underlies performance of more complex intelligence tasks, the first set of training studies investigated whether training these basic processes improves abilities that rely on them.

One of the first articles taking on this issue was published by Klingberg, Fornsberg, & Westerberg [28]. They trained children with attention-deficit/hyperactivity disorder (ADHD) by means of WM training. WM deficits seem to play a central role in children with ADHD, which was the rationale for choosing children with this particular disorder. By directly training these deficits, the authors hoped not only to improve overall ADHD symptoms, but also to obtain some general improvements in cognitive performance. Klingberg et al. used a training regimen consisting of several computerized tasks: (1) a visual span task where circles appeared one at a time in different locations of a four by four grid. Participants were instructed to indicate the positions of the circles in the correct order; (2) a backwards digit-span task where participants were required to repeat a spoken series of digits in the reverse order; (3) a letter-span task where participants were presented with a series of letters. After the presentation they were probed for a certain position in the series and were required to reproduce the letter at that position. Finally, (4) a go-no go task in which two grey circles were presented on a computer screen. As soon as one of the two circles changed color from grey to green, participants were required to press a corresponding key as quickly as possible. However, if the circle color changed from grey to red, no key press was required. A very important characteristic of the training intervention of Klingberg et al. was its adaptivity, that is, the task increased in difficulty as the participants’ performance improved, or decreased in difficulty if performance decreased. In Experiment 1, Klingberg and colleagues compared 7 children in a training group with 7 children in a no-contact control group. The mean age of the participants was eleven years and all were diagnosed with ADHD. Participants trained for 25 minutes per session for 5–6 weeks. Before and after the training, Raven’s Colored Progressive Matrices (CPM) [29] were administered among other measures. The authors reported a significant performance improvement in the trained group compared to the controls. In Experiment 2, Klingberg et al. trained four young adults with the same training regimen. They used a more difficult version of Raven’s matrices, the Advanced Progressive Matrices (APM) [12]. Again, the authors report a significant performance increase in intelligence after training. It must be noted, however, that the generalizability of these results is limited because only four subjects were trained and no matched control group was used.

In a larger follow-up study, Klingberg et al. [30] used a commercially available training program (RoboMemo®, Cogmed Cognitive Medical Systems AB, Stockholm, Sweden), which is very similar to the intervention described in the previous study. Klingberg and colleagues tested 20 participants in the experimental condition and 24 participants in the active control condition in a randomized-controlled trial. The average participant age was ten years and all were again diagnosed with ADHD. Training time was 40 minutes per session for a time span of five weeks. Again, Raven’s CPM was used to assess intelligence. The authors reported a significant gain in intelligence and also persisting benefits in the experimental group (as compared to the control group) even three months after training completion.

Unfortunately, the findings reported above have not easily been replicated. Klingberg’s group for example conducted another study with healthy preschool children using the same training regimen as used in the 2005 study [31]. In contrast to the earlier studies, there was no improvement in intelligence as measured by the Wechsler Preschool and Primary Scale of Intelligence – Revised (WPPSI-R) [32], although they found benefits to other untrained tasks. Similarly, Holmes, Gathercole, & Dunning [33] were not able to replicate improvements in intelligence of children with poor WM performance by using the training program of Klingberg et al.

Our own work with healthy young adults [34] is heavily based on the assumption that WM and intelligence share a common capacity constraint [35]. Thus, our rationale is that if we train participants on one domain, we should see benefits on other domains which share similar processes. In addition, we propose that a successful training task must fulfill several criteria in order to be successful: (1) the task should minimize the development of task specific strategies; (2) the training must be adaptive to allow participants to train at the peak of their performance in order to prevent automatization on one hand, and excessive demands on the other hand; (3) the task should be complex enough to train several different processes at once in order to maximize process overlap with other tasks. One task that seems to fulfill these criteria is a dual n-back task as we had used previously [36, 37].

In this task, participants are presented with a stream of stimuli one after another. Each time the current stimulus is the same as the one presented n stimuli back in the stream, the participant is required to press a key. Since we used this task as a dual task, participants were required to perform this task simultaneously in a visual and an auditory modality.

The dual n-back task is available as part of the software "Brain Twister" available from the University of Bern which is provided free of charge if used for research [38].
In order to obtain adaptivity, the level of n was adjusted according to the actual performance of the participant. A total of 34 participants were tested and trained and then compared on performance with 35 no-contact controls. Participants were 26 years old on average. The training intervention lasted either 8 days, 12 days, 17 days, or 19 days; training time per day was approximately 25 minutes. We used either Raven’s APM or the Boehmner Matrices Test (BOMAT) [39] to assess improvements on intelligence. We found a significant group (training vs control) by session (pre vs post) interaction, showing that the training group improved more than the control group overall. Moreover, we were also able to show a dose-response curve, that is, with increased training time there was increasingly greater improvement in matrix reasoning.

Very recently, we were also able to show that single n-back training is equally effective as dual n-back training [40]. In this study, we trained a total of 47 participants either on a single n-back task or on a dual n-back task. Performance of those training groups was compared to a no-contact control group consisting of 43 participants. The average age was 19 years. The training lasted over a span of four weeks with approximately 20 minutes of training per day. We used two different matrix reasoning tests, the Raven’s APM and the BOMAT in order to assess intelligence. We found a significant group (training vs control) by session (pre vs post) interaction for both intelligence tasks and in both training groups. With this study, we could not only replicate our previous findings [34], but also show that a considerably less complex task, a single n-back task, is as equally effective as a dual n-back task.

Turning to more executive approaches, Rueda, Rothbart, McCandliss, Saccomanno, & Posner [41] tested the impact of attention training on intelligence in young children. Their training paradigm consisted of a series of nine (Experiments 1 and 2) or ten (Experiment 3) different tasks that are related to executive attention and were presented in a game-like fashion. Each task consisted of several levels of difficulty and became more challenging as participants’ performance improved. 24 four-year-old children and 12 six-year-old children were trained. Training was conducted on five days over a 2–3 week period. In order to assess intelligence, the Kaufman Brief Intelligence Test [42] was conducted before and after the intervention. The authors reported a significant group by session interaction in the four-year-old group, indicating that the experimental group significantly improved intelligence performance compared to an untrained control group. This effect was not observed in the group of six-year-olds. Rueda and colleagues explain this effect by assuming that intelligence-related tasks and tasks that require the neural executive attention network recruit similar brain regions.

It has been argued that neural plasticity is more prevalent in younger years than in later life [cf. 43]; thus, the chances to improve intelligence are assumed to be higher in younger than older people. Although we provided evidence that intelligence can be improved in young adults [34, 40], Garlick [43] assumed that intelligence can only be altered in people younger than approximately 16 years, because the neural connections are not as fixed as in later years. In a very recent study, Karbach & Kray [44] investigated this issue with a training study involving children, young adults, and old adults. They investigated the effect of task-switching training on Raven’s Standard Progressive Matrices (SPM) [45] as well as figural reasoning [cf. 46]. Their participants trained with a task-switching paradigm in which they were required to alternate between two different tasks every other trial. Task A required a category response towards a presented picture (e.g., is the picture a tree or a flower), and Task B required an attribute judgment (e.g., size or color) towards a presented picture. Karbach and Kray trained a total of 126 participants, with an active control group consisting of 42 participants, equally distributed among three age groups. The mean age of the age groups was 9, 22, and 69 years respectively. Subjects trained for four days, 30–40 minutes per day. In contrast to the plasticity hypothesis, Karbach and Kray found significant intelligence improvements in all three of the trained age groups. They discuss their findings in that their intervention involved the training of several processes, such as goal maintenance or interference resolution which overlap with the processes required to perform the intelligence tasks.

### Improving intelligence by other interventions

Besides interventions that are based on training WM and executive functions, there are also other methods that demonstrated an improvement in intelligence. This group of studies contains a wide variety of approaches towards improving intelligence.

For example, Basak and colleagues [47] took an off-the-shelf real-time strategy video game as their training vehicle (Rise of Nations: Gold Edition developed by Big Huge Games and published by Microsoft Game Studios in 2004). Basak et al. argued that this type of training predominantly requires the ability to switch between different goals and maintaining multiple items in WM, which, if trained, might lead to improvements in tasks that were not part of the training. They trained and tested 19 old adults with the game and compared this group to a no-contact control group consist-
sisting of a selection of different tasks that assessed concepts such as processing speed, WM, inductive reasoning, visual-spatial processing, and divergent thinking. They tested and trained 87 participants in the experimental group and tested 63 participants in a no-contact control group. The average age of the subjects was 73 years. The authors reported a significant improvement in fluid abilities in the experimental group after the intervention.

Schellenberg [51] investigated the impact of music lessons on a full-scale IQ test. Schellenberg's training rationale was based on the assumption that experiences related to music training, such as focusing attention, memorization of music passages, reading music notation, or mastery of fine-motor skills would have a positive impact on cognitive ability in general. In line with Garlick's reasoning [43], he also argued that this is especially true for children, whose brain structures have higher plasticity than those of older persons [c.f. 43]. In order to test this hypothesis, a total of 144 children were randomly assigned to one of four different groups: keyboard lessons, voice lessons, drama lessons, or no lessons. The average age of the children was six years. The lessons were taught for 36 weeks and the children were pre- and post-tested on the Wechsler Intelligence Scale for Children-Third Edition (WISC-III) [52]. After collapsing the keyboard lesson and voice lesson groups into an experimental group and the two other groups into a control group, Schellenberg found a reliably larger increase in intelligence scores in the music training group. The average increase for the experimental group was 7.0 IQ points while controls increased on average only 4.3 IQ points. Based on these results, Schellenberg concluded that music lessons have a small but nevertheless reliable positive impact on intelligence.

There have been some attempts to increase intelligence by pharmacological means, with mixed evidence thus far. Although there have been a few studies with humans demonstrating that certain psychomotor stimulants and D2 dopamine-receptor agonists seem to have some effects on isolated cognitive processes [53, 54], there is only one study, to our knowledge, that showed a positive impact on intelligence. Rae, Digney, McEwan, & Bates [55] used an oral creatine supplement to provide the brain with additional energy in order to prevent limited energy resources in the brain due to heavy cognitive workload. In their double-blind, placebo-controlled, crossover design, Rae and colleagues supplemented their vegan and vegetarian participants with daily 5 g of creatine monohydrate over a period of 6 weeks, followed by a 6 week wash-out period, and again followed by a supplemented period of 6 weeks. The vegan/vegetarian participants were selected because creatine levels are lower in these people than in omnivores. They tested 45 subjects with an average age of 26 years on Raven's APM at the beginning of the study and at the end of each supplemented period, and finally, at the end of the wash-out pe-
Improving intelligence

period. The statistical analysis showed a significant improvement after the supplemented period compared to after the wash-out period. Unfortunately, the generalizability of this study is limited since it

is unknown whether similar effects could be obtained in omnivores with normal baseline levels of creatine.

Conclusions and future directions

We reviewed 11 studies that provide accumulating evidence that performance in intelligence tests can be positively altered by some form of intervention. The reviewed studies are heterogeneous on several dimensions, one of which being the theoretical rationale. The largest group of studies is currently the one that tries to improve intelligence by training of WM or executive processes [28, 30, 34, 40, 41, 44]. The rationale of this group of studies is based on a large body of research showing the close relationship between WM and intelligence. Other approaches investigate the impact of video games on intelligence [47], or apply cognitively stimulating activities to improve intelligence [48, 50]. Although the latter two approaches seem very distinct on first sight, their outcomes may be driven by similar mechanisms. For example, both studies challenge older participants to perform problem-solving activities, as explicitly induced by Tranter & Koutstaal [48] and by Stine-Morrow, Parisi, Morrow, & Park [50], or more implicitly induced by Basak, Boot, Voss, & Kramer [47] who asked their participant to engage in a real-time strategy game that required problem solving skills as well. By providing children with music lessons and thereby exposing them to new experiences that are unique to musical instructions, Schellenberg [51] successfully improved full-scale intelligence performance. Finally, Rae, Digney, McEwan, & Bates [55] were also able to improve intelligence performance by supplementing participants with creatine, assuming that this supplement gives additional energy support to the brain when energy demands are high.

Although each of the approaches referenced above have a reasonable theoretical rationale, it is far from clear how and why these interventions work [see also 56]. An important step to shed light on this issue has recently been made by Dahlin, Neely, Larsson, Backman, & Nyberg [57], who showed that a cognitive training intervention led to improvements in an untrained task only if these two tasks activate similar regions in the striatum. Furthermore, McNab and colleagues [58] found that the cortical dopamine D1 receptor binding potential changes in prefrontal and parietal brain regions after WM training. Studies like these are important for further understanding of the mechanisms of training and the resulting benefits in intelligence.

Unfortunately, many of the reviewed studies have certain methodological shortcomings that can be criticized. For example, the repeated usage of the identical test material within a study in order to measure intelligence is unfortunate, because practice on the same items considerably lowers the sensitivity to assess intelligence processes [e.g., 16, 17]. Another shortcoming is the inclusion of a no-contact control group. Although the inclusion of such a control group might be appropriate to control for re-test effects and in order to show that the investigated intervention works in principle [59], future research should include active control groups in order to control for un-specific effects like for example the time of engagement with the computer, social interaction, or more importantly, motivational effects. Also, until now, only a very restricted range of intelligence tasks has been used within each study, and there is little evidence that the effects go beyond laboratory tasks to standardized measures or even academic achievement or into daily life in general. Finally, very few studies have looked into the long-term effects of their interventions; thus, it is not known how long the improvements last or whether some “booster-sessions” might be beneficial in order to maintain performance [c.f. 14, 56, 60].

In our opinion, an ideal training study should incorporate the following features: the implementation of (1) a significant number of participants; (2) a randomized assignment of the participants either to the experimental or the control group in order to control for underlying motivational differences; (3) an active control group that engages in activities that are as similar as possible to the experimental group with the exception of the portion that is supposed to be responsible for the improvement in intelligence; (4) a careful selection of multiple tests that assess different aspects of intelligence and that can be reliably used to assess intelligence on more than one single occasion; and finally (5), an assessment of long-term effects in all groups, not just the experimental group.

Although it is easy to conceptualize the ideal study, one has to keep in mind that the logistics for running an intervention study are usually very challenging, expensive, and time consuming. Therefore, progress may not take place as quickly in this field as it does in others. But, as also pointed out by Sternberg [22], we believe that a first important step has been taken by showing that human abilities are not as static as was previously hypothesized [e.g., 18]. Nevertheless, we would like to add a word of caution that the reviewed field of research is relatively young and first successful attempts to improve intelligence have only emerged in the last couple of years. Despite this limited evidence, the availability of cognitive training inter-
ventions has dramatically increased, mainly because they are now readily available for personal computers and also for popular mobile devices. These commercially available training interventions are often advertised as being highly effective in improving general mental capacity, but unfortunately, scientific proof for such claims is rarely provided [61].

Further work must be done to uncover the underlying mechanisms that promote benefits to intelligence and to investigate how meaningful the effects are in real life [56, 62]. Therefore, future studies should broaden their measures of intelligence, as well as incorporate direct measures of real-world performance such as on-the-job or academic achievement. Furthermore, it is important to have intervention strategies at hand that are applicable to a broad range of people independent of age and health status. Finally, it is also of great interest whether the gain on intelligence depends on training features such as spaced or massed training [e.g., 63] and how interindividual differences mediate the effects of training.

References

8 Rohde TE, Thompson LA. Predicting academic achievement with cognitive ability. Intelligence. 2007;35(1):83–92.
25 Kyllonen PC, Christal RE. Reasoning ability is (little more than) working-memory capacity? Intelligence. 1996;14:389–413.

Correspondence:
Martin Buschkuehl
Department of Psychology
The University of Michigan
East Hall
530 Church Street
Ann Arbor, MI, 48109-1043
E-Mail: mbu@umich.edu
Improving intelligence

49 Cartell RB, Cartell AKS. Handbook for the individual or group Culture Fair Intelligence test. Champaign: IPAT; 1960.