When (deliberate) practice is not enough – the role of intelligence, practice, and knowledge in academic performance

Merim Bilalić[1] · Ratko Đokić[2,3] · Maida Koso-Drljević[2] · Nermin Đapo[2] · Thomas Pollet[1]

Abstract

It is well established that academic performance (AP) depends on a number of factors, such as intellectual capacities, practice, and previous knowledge. We know little about how these factors interact as they are rarely measured simultaneously. Here we present mediated-Factors of Academic Performance (m-FAP) model, which simultaneously assesses direct and indirect, mediated, effects on AP. In a semester-long study with 118 first-year college students, we show that intelligence and working memory only indirectly influenced AP on a familiar, less challenging college course (Introduction to Psychology). Their influence was mediated through previous knowledge and self-regulated learning activities akin to deliberate practice. In a novel and more challenging course (Statistics in Psychology), intellectual capacities influenced performance both directly and indirectly through previous knowledge. The influence of deliberate practice, however, was considerably weaker in the novel course. The amount of time and effort that the students spent on the more difficult course could not offset the advantage of their more intelligent and more knowledgeable peers. The m–FAP model explains previous contradictory results by providing a framework for understanding the extent and limitations of individual factors in AP, which depend not only on each other, but also on the learning context.

Introduction

Most people would agree that a grade on a college course depends on the intellectual capacities and the effort of the student. Some would also argue that previous knowledge can play an important role. Yet, empirical studies have repeatedly shown how difficult it is to pinpoint the exact contributions of intellectual capacities, practice, and previous knowledge in academic performance. Intellectual

Merim Bilalić and Ratko Đokić These authors share first authorship.

✉ Merim Bilalić
  merim.bilalic@northumbria.ac.uk

[1] Northumbria University, Ellison Square,
   Newcastle NE1 8ST, UK

[2] Department of Psychology, Faculty of Philosophy,
   Sarajevo University, Franje Rackog 1, 71000 Sarajevo,
   Bosnia and Herzegovina

[3] Present Address: Institute for Social Research in Zagreb
   (IDIZ), Frankopanska 22/I, 10000 Zagreb, Croatia

capacities, commonly measured by intelligence tests, lose their impact on academic performance as pupils move from middle school to high school (Neisser et al., 1996), and then on to college (Jensen, 1998; Poropat, 2009; Wachs & Harris, 1986). Practice, measured as effort-related activities, is often not a significant predictor of academic performance at all (Rau & Durand, 2000; Schuman et al., 1985). Only specific practice activities tend to be relevant predictors of academic performance (Nandagopal & Ericsson, 2012a). The impact of previous knowledge also varies greatly, from nonexistent (Federici & Schuerger, 1976; Griggs & Jackson, 1988) to high (Bloom, 1976; Sadler & Tai, 2001; Tai et al., 2005). Here, we demonstrate that the previous ambiguous results are related to two factors: a) the lack of simultaneous measurement of all important factors, which precludes uncovering indirect influences; and b) the requirements and difficulty related to the novelty of the college course, both often neglected in research, which may accentuate different factors.

Factors of academic performance

Intellectual Capacities

Traditionally, intellectual capacities are considered to be the main predictor of academic performance (Horn & McArdle, 2007; Spearman, 1904). This is evident in the common practice of using the Scholastic Aptitude/ Assessment Test (SAT) and Graduate Record Examination (GRE), which are correlated with psychometric tests of intellectual capacities (Frey & Detterman, 2004), for selection purposes. This makes sense in as much as intelligence, broadly defined as mental power (Spearman, 1904), should enable an individual to grasp connections within the new material and allow for quick orientation in new situations (for a more incremental and malleable definition of intelligence, see Yeager & Dweck, 2020). Students who score higher on intelligence tests should learn new material better and more quickly, which in turn should result in better academic performance, or intellectual performance more generally, all other factors being equal (Vaci et al., 2019b). Intelligence is indeed a good predictor of success in middle and high school (Gagné & St Père, 2001; Neisser et al., 1996).

Its predictive power, however, weakens at college level, where it is either a smaller factor (Jensen, 1998; Poropat, 2009; Roberts et al., 1978) or not a significant contributor to academic success in any way (Wachs & Harris, 1986). Unlike primary and secondary education, which most people will complete, only a select cohort, often based on intelligence as much on their previous academic success, will go to college. This filtering process inevitably leaves highly intelligent students in the pool of college students. The restriction of range on a measure is known to lead to a reduction in the predictive power of that measure (Pearson, 1902; Vaci et al., 2014). However, even when we account for the restricted range, reviews demonstrate that intelligence explains only one quarter of the variance in academic performance (Neisser et al., 1996; but see, Zaboski et al., 2018). It is therefore possible that other factors, such as different developmental stages, play a role in the diminishing returns of intellectual capacities at the college level. For example, social expectations and individual needs, also known as developmental tasks (Havighurst, 1953), are vastly different during childhood, adolescence, and early adulthood, which may impact people's academic interests and performance (Masten & Coatsworth, 1998; Tetzner et al., 2017).

Most of the research on the role of intellectual capacities on academic performance employed classical intelligence tests (e.g., Debatin et al., 2019; Primi et al., 2010; Soares et al., 2015). The relation between intelligence tests and other ability tests, most notably working memory tests, is highly debated (e.g., Blair, 2006; Heitz et al., 2006). Depending on the view, captured abilities by working memory tests are either a part of larger abilities captured by intelligence tests (e.g., Kyllonen, 2002; Stauffer et al., 1996) or an independent entity (e.g., Ackerman et al., 2005; Engle, 2018). What is less controversial is that abilities such as early retention and later recollection are captured rather by working memory tests than by intelligence tests (Unsworth, 2016; Unsworth & Engle, 2007). Given that exactly these abilities are crucial for academic performance (Barrouillet & Lépine, 2005; Bergman Nutley & Söderqvist, 2017; Gathercole et al., 2006) working memory has been extensively studied as a predictor of academic attainment, mainly in the sphere of reading and mathematics in younger (preschool, elementary or high school) children (for a review, see Peng et al., 2016; Swanson & Alloway, 2012). However, less emphasis has been given to the relationship between working memory and science learning at the college level (Yuan et al., 2006).

Previous Knowledge

One of the possible reasons for the diminished predictive power of intelligence at the college level might be the knowledge that students bring to courses (Dochy et al., 2002; Shapiro, 2004). College courses often build on previous knowledge. For example, the course "Statistics for Psychologists" is often built on probability calculus (i.e. combinatorics), a field of mathematics which is usually taught at high school. The existence of relevant knowledge in long-term memory allows for an easier understanding of new related material (Chase & Simon, 1973; Gobet & Simon, 1996). The new incoming information will be better understood, better remembered, and finally, better recalled if there is an already existing knowledge base where the information can be integrated (Bilalić, 2017; Hambrick, 2003; Hambrick et al., 2008). Consequently, previous knowledge is often (Diseth, 2011; Diseth et al., 2010; Dochy et al., 1999), but not always (Federici & Schuerger, 1976; Griggs & Jackson, 1988), a significant factor in predicting academic performance. Its impact arguably reflects the degree of overlap between existing knowledge and the new knowledge that students need to acquire (Nandagopal & Ericsson, 2012b).

It is plausible to assume that more intelligent students come with a larger knowledge base related to the course. The existing knowledge may then water down the influence of intelligence on academic performance as knowledge may already explain part of the performance that would otherwise be explained by intellectual capacities. Even in rare cases where less intellectually capable individuals acquired more knowledge, the consequence would be the same – the

relation between intellectual capacities and academic performance would be weakened.

Practice

Besides intelligence and knowledge, how much time one spends studying the course material is often considered an important predictor (Nandagopal & Ericsson, 2012b). Although the link between course-related activities and course performance seems self-evident, it turns out that it is difficult to find studies which have established such an association. Often, the amount of time spent on learning is not significantly related to performance (Macnamara et al., 2014; Schuman et al., 1985), or the relationship may even be a negative one (Chinn et al., 2010; Nonis & Hudson, 2006). Only when the *quality* of learning, as well as the quantity, is accounted for, does practice become a significant predictor (Nandagopal & Ericsson, 2012a; Plant et al., 2005; Richardson et al., 2012). Students who actively participate in their own learning process by setting goals and reflecting on the effectiveness of their learning (metacognition) are more successful than their peers who plan less and are less goal-driven (Zimmerman, 2008).

These self-regulated learning strategies (see also, Deslauriers et al., 2019) are similar to focused deliberate practice found in experts in different domains (Ericsson et al., 1993; Nandagopal & Ericsson, 2012a). Experts' engagement in activities which provide most opportunities for improvement of performance also requires motivational and metacognitive components similar to self-regulated learning strategies (Nandagopal & Ericsson, 2012a). Arguably the biggest difference between the expert and academic contexts is the presence of feedback from a teacher/coach in the former and its absence in the latter. The immediate feedback from an instructor also happens to be a controversial aspect of the deliberate practice definition (Ericsson & Harwell, 2019; Ericsson, 2020, 2021; Macnamara & Hambrick, 2021). This can be seen by a number of studies which measured practice activities without immediate feedback from an instructor but were classified as measuring deliberate practice (Macnamara et al., 2014, 2016), especially in the education setting (Ericsson & Harwell, 2019).

Mediated – Factors of Academic Performance (m-FAP) model

One problem with previous research is that there are hardly any studies on academic performance that simultaneously measure all three relevant constructs. Usually researchers focus on a single factor, for example, investigating the influence of intellectual capacities (Debatin et al., 2019; Primi

et al., 2010; Soares et al., 2015) or practice (Doumen et al., 2014; Nonis & Hudson, 2006; Rosário et al., 2013). It is conceivable that the focus on a single factor could lead to wrong conclusions. For example, an apparent lack of association between intellectual capacities with academic success may be moderated by widely differing effort and investment among students.

When more factors are taken into account, often the interplay between them is not examined (Castejón et al., 2006; Diseth et al., 2010; Limanond et al., 2011). It is therefore difficult to pinpoint the actual mechanism of academic performance. One does not know whether academic performance is a consequence of, for example, practice and effort independent of previous knowledge and intelligence. Similarly, we do not know to what extent these factors cause academic performance. This is unfortunate, as many researchers have argued that intelligence, knowledge, and practice are interconnected in many performance-based contexts (Burgoyne et al., 2019a, b; Hambrick et al., 2018; Mosing et al., 2019; Vaci et al., 2019a, b), including academic performance (Dochy et al., 1999).

Here we propose a new theoretical model to simultaneously investigate the interplay between the main factors of academic performance. We call this new framework the mediated-Factors of Academic Performance (m-FAP) model, because it is based not only on the theoretically assumed direct relations between the concepts, but also on indirect effect, that is mediations (see Fig. 1 for theoretical connections between intelligence, knowledge and practice). For example, students generally develop interests before they enter college education and pursue those interests in a formal college setting. This leads to exposure to information and the acquisition of relevant knowledge. It is a fair assumption that more intelligent individuals, who also tend to be more
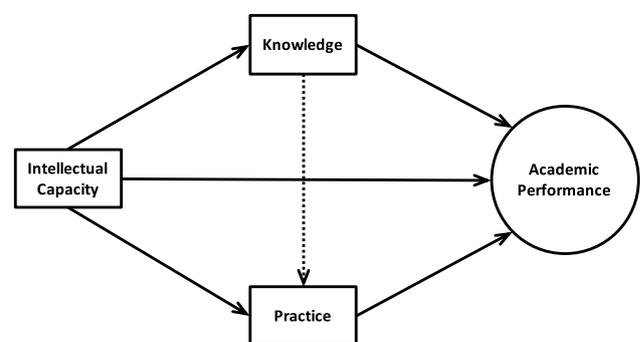
Fig. 1 Mediated-Factors of Academic Performance (m-FAP). All three factors, intelligence, previous knowledge, and practice, may influence academic performance directly. Intelligence, however, may influence performance indirectly, through both previous knowledge and practice. Previous knowledge, for example, could also influence academic performance through practice

intellectually curious (von Stumm et al., 2011), will also profit more from extended exposure and gain more relevant knowledge. Once in college, more intelligent students may generally need less time to acquire the same amount of necessary material. Consequently, intelligence may influence academic performance, not only directly, but also indirectly through previous knowledge and practice.

Intellectual capacities may also influence academic performance through practice. More intelligent students may need to spend less time on the material to achieve similar results as their less able peers (Deary, 2011; Neisser et al., 1996; Spearman, 1927; Vaci et al., 2019a, b). A similar interplay between previous knowledge and practice may be expected. Students with more relevant knowledge will need less time to understand new information, which may lead them to spend less time on the material (Hambrick, 2003; Hambrick et al., 2008).

Current study

Previous studies often also fail to take into account the requirements and difficulty related to the novelty of individual courses. Academic performance is often an average measure across all courses, which is suboptimal as the relations between the three factors may be different depending on the requirements of a particular course. A familiar and consequently a relatively easy course, where only facts and the relations between them are acquired, may not be particularly suitable to establish whether intelligence has a direct influence. Less intelligent students may do equally well, or maybe even better, by working hard. In contrast, novel, more challenging courses, which require assimilation of concepts that are not straightforward and extensive use of logical thinking, may not reward practice and hard work to anywhere near the same extent. Here, intellectual faculties and relevant previous knowledge may be of more use (Spearman, 1927). The situation is arguably even more accentuated in short semester-long courses where there is simply no time to compensate for the lack of knowledge and lesser intellect through sheer effort.

Our current study seeks to address the gap in the literature by not only simultaneously measuring relevant factors and their interplay with our m-FAP model, but also examining their predictive power on a traditionally novel and difficult college course and a more familiar, easy course. We tested an undergraduate cohort of 118 psychology students, who in their very first semester of college education attended two single-semester-long courses: A) Introduction to Psychology, a generally straightforward fact-based course, which was highly familiar to students due to their interest in the topic and their previously acquired high-school knowledge; and B) Statistics, traditionally the bogeyman of psychology students, a novel topic outside of students' interest, where logical thinking and understanding of probabilities is necessary.

We measured students' performance in multiple exams throughout the course of the semester, as well as their intellectual capacities, previous knowledge, and amount and quality of practice (see Method and Supplementary Material, SM). Intellectual capacities were measured by standard intelligence tests as well as working memory (WM) tests. Previous knowledge was operationalized with separate tests of psychological and mathematical knowledge at the beginning of the semester. Practice was assessed regularly throughout the semester with weekly questionnaires and was validated with individual diaries. We distinguished between goal-directed, self-regulated solitary practice and all other learning activities. We call the solitary learning activities deliberate practice (DP), leaning on the terminology from the expert performance framework (Ericsson & Harwell, 2019; Ericsson et al., 1993; Ericsson, 2020; Macnamara & Maitra, 2019; Macnamara et al., 2014). We call other learning activities, which did not involve planning and often involved various distractors (e.g. TV, radio, internet), non-deliberate practice (NDP). We supplement the differentiation between the two kinds of practices by the Motivated Strategies for Learning Questionnaire – MSLQ (Pintrich, 1991), which measures learning strategies used by the students when dealing with the course material.

One of the advantages of the m-FAP model is that it can easily accommodate new factors, such as the course context here. We expect intellectual capacities to have a direct impact, but also to indirectly influence academic performance through knowledge and (deliberate) practice. For the familiar and less challenging course, Introduction to Psychology, we expect a comparatively weak direct relation of intellectual capacities with success. Intellectual capacities will exert most of their influence on academic performance through previously accumulated knowledge and (deliberate) practice. In contrast, for a novel and consequently a more challenging course, Statistics, we hypothesize that intellectual capacities will still directly influence performance, on top of their indirect influence through knowledge and practice. Overall, intellectual capacities should have more impact in the more challenging context. Practice in turn will have a weaker influence in the novel and challenging course than in the less challenging one because dependence on previous knowledge will render it less effective.

We also expect intelligence and WM measures to differ in their explanatory power depending on the course context. In the Introduction to Psychology course, where the memorization of information and facts is necessary, WM will be of more relevance than intelligence. In contrast, the Statistics

course, which requires understanding of statistical concepts, will rely more heavily on intelligence than on WM.

Method

Participants

The participants were the population of the first year of undergraduate studies in Psychology, altogether 139 students, as all students agreed to take part in the study. We excluded 19 students who were taking the year again as they had already listened to the lectures (they also did not participate in the Introduction to Psychology course). Two further students were excluded because they never took any exams in Statistics, leaving 118 students ($M_{age} = 19.1$, $SD = 1.4$, range $= 17 - 27$ years, 90% women). All of them had taken exams in both Introduction to Psychology and Statistics, but some were not present when other measures were taken: 3 students, or 2%, missed intelligence tests; 11, or 9%, missed working memory tests; 6, or 5%, missed practice estimates for Introduction to Psychology; 4, or 3%, missed knowledge tests for Statistics; and 21, or 18%, missed knowledge tests for Introduction to Psychology. We consider these missing values to be random (see SM, Sect. 1) and have analyzed the data using standard imputation techniques (Van Buuren, 2018). It should be noted that the analyses without imputation produced essentially the same pattern of results (see SM, Sect. 6). All participants gave written informed consent that their data can be used for research purposes and published in anonymous form. The research protocol was approved by the internal ethics committee of the Department of Psychology at the University of Sarajevo. The students were debriefed at the end of the semester and were allowed insight into their individual results if requested.

Power Analysis

The average influence of intelligence on academic performance (Intelligence $\rightarrow$ AP) is around (standardized) $\beta = 0.18$ (Poropat, 2009, p. 329, Table 2; Richardson et al., 2012, p. 366, Table 6). We took the lower effect estimate (20% of the range) for Introduction to Psychology ($\beta = 0.03$), and the upper effect estimate (80%) of the average effect (Richardson et al., 2012) for Statistics ($\beta = 0.29$). We expect a strong relation, around $\beta = 0.50$, between previous knowledge and intelligence (Intelligence $\rightarrow$ Knowledge) for both courses (Hambrick et al., 2008, p. 273, Table 8). Previous knowledge should have a strong association with performance on both courses (Knowledge $\rightarrow$ AP), but should be larger in Statistics than in Introduction to Psychology. Given that the association

of previous knowledge/achievement with academic performance is between $\beta = 0.32$ and 0.36 (Schneider & Preckel, 2017, p. 26), we decided to take the associations of 0.30 and 0.50 for Introduction to Psychology and Statistics, respectively. This is arguably an underestimate, as our knowledge tests were directly relevant to the academic performance (Bloom, 1976; Dochy et al., 2002), unlike most previous research, which used the measures of previous success as an indicator of knowledge/achievement (Schneider & Preckel, 2017). Similarly, we used specific college courses, instead of the average across a wide range of courses.

Based on these considerations, the necessary number of participants for discovering the mediation between intelligence and previous knowledge on academic performance with a power of 0.80 and the alpha level of 0.05 is around 106 for Introduction to Psychology and 107 participants for Statistics (Zhang, 2014). The 0.80 power with the 0.05 alpha level for discovering the total effect of intelligence on academic performance, which includes both the direct (Intelligence $\rightarrow$ AP) and the mediation effects (Intelligence $\rightarrow$ Knowledge $\rightarrow$ AP), is around 262 for Introduction to Psychology and 75 participants for Statistics. Most importantly, to uncover the difference between these two total effects of intelligence on academic performance with a 0.80 power, one would need around 94 participants in both groups (see SM, Sect. 2, for the detailed calculations). We did not calculate power analysis for the WM (via Knowledge and DP) as we are not aware of the reliable estimates for the necessary relations.

Procedure

Table 1 shows the overall schedule during the first semester (16 weeks). IQ measures were administered in two sessions (first Raven, then VGAT and NGAT), in groups of up to 60 participants in weeks 2 and 3 of the semester. WM measures were administered during a single session, in groups of four participants, in weeks 5 and 6. The knowledge tests were administered in the first week. The exams were given in week 8 (both courses) and week 16 (both courses). There was an additional exam for the statistics course in week 15. The eight questionnaires and three diaries were administered for Introduction to Psychology, thereby covering 70 days (65%) of a total of 108 days in the semester. The questionnaires and diaries overlapped in one week. For Statistics, nine questionnaires and four diaries were administered, covering 82 days (76%) in the semester. The questionnaires and diaries overlapped in two weeks. Students filled in the questionnaires at the beginning of the regularly scheduled workshop classes, retrospectively for the last seven days. Students filled the diaries at their homes, concurrently for a given period, and handed them in on the first day after completion.

Table 1 Study overview. The time of administration of Knowledge, Intelligence, and WM tests, Motivated Strategies for Learning Questionnaire, exams and practice activities for Introduction to Psychology and Statistics

	WEEK OF SEMETER															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
INTRODUCTION TO PSYCHOLOGY			Q	Q	Q	Q+D	D			Q	Q		Q	Q	D	
	K	IQ	IQ		WM	WM		E1		MSLQ						
STATISTICS		Q	Q	Q	Q	Q+D	Q						Q+D	Q	D	
	K	IQ	IQ		WM	WM		E1		MSLQ					E2	E3

Note. K = Knowledge Test, Q = Activity Questionnaire, D = Activity Diary, E = Exam, IQ = Intelligence Tests, WM = Working Memory Tests, MSLQ = Motivated Strategies for Learning Questionnaire

Intelligence (Intelligence) tests

Three intelligence tests were administered: Raven Advanced Progressive Matrices (Raven), Verbal General Ability Test (VGAT), and Numerical General Ability Test (NGAT). The Raven measures abstract reasoning (Raven, 1991), the VGAT measures verbal ability (Smith, 1999), and the NGAT measures numerical ability (Smith, 1999). The reliability, as measured by Cronbach's alpha, was 0.77 for Raven, $\alpha = 0.77$ for the VGAT, and $\alpha = 0.81$ for the NGAT.

Working Memory (WM) measures

Three WM measures were measured: operational span (OSPAN), reading span (RSPAN), and symmetry span (SSPAN). All three measures capture the WM or the participants' ability to remember and correctly recall material while they were working on a different task (Conway et al., 2005; Đokić et al., 2018; Redick et al., 2012). The reliability in this particular case were: $\alpha = 0.81$ for the OSPAN, $\alpha = 0.84$ for the RSPAN, $\alpha = 0.72$ for the SSPAN.

Practice measures

The practice questionnaire and diary were designed for the collection of information about specific learning activities, duration, and circumstances of individual episodes of learning. The goal was to distinguish between focused self-regulated learning episodes, which we call deliberate practice (DP), and other, less efficient activities (called non-deliberate practice, NDP). In devising the questionnaire and diary (see SM, Sect. 3, for the actual instrument) we followed recommendations from the research on survey methodology (Schwarz & Sudman, 2012) and autobiographical memory (Conway et al., 2005). Firstly, we kept the reference period for which the students were asked about their activities (one week) as short as possible but at the same time without too frequent questioning (Bradburn et al., 1987). Secondly, in order to avoid "telescoping" (allocating the activities in question earlier or later then they really happened)

the practice inquiry was aided with "landmarks", significant general dates, as the retrieval cues (Loftus & Marburger, 1983; Shum, 1998; Sudman & Bradburn, 1982). The questionnaire was structured around individual learning episodes in every day of the previous week. It asked about 1) duration, 2) specific activity (e.g. theory, exercises), 3) information about individual or group work, and 4) whether other distractors were presents (TV, radio, internet, interruption by other persons) during the individual episodes. The students also judged the quality of the work. This structure provides retrieval cues and clarifies the questions, as well as offering a framework for response.

The diary featured essentially the same structure where the individuals were registering their individual learning episodes. The format was the same as in the questionnaire (e.g., duration, activity, individual/group, distraction/no-distraction) with one exception: it did not ask about the quality of the activity, but rather about the difficulty and effort. Both difficulty and effort were measured with a rating scale which went from 1 (lowest) to 7 (highest).

The practice diaries and questionnaires also featured questions about individual instructions, which would come closest to the traditional definition of the DP. However, only a handful of students (14 or 6% across both courses) were indicated having such instructions for only a short period of time (less than a minute per week when averaged over the whole period). We consequently used solitary learning without distractions for the DP. This kind of learning indicates advanced planning and has been used as an indicator of deliberate practice and self-regulated learning in previous studies (Nandagopal & Ericsson, 2012a; Plant et al., 2005). All other learning activities were not considered to be DP and were assigned to the NDP variable. The values in the DP and NDP variables were the estimates from the questionnaire and the diary. Where the diary and questionnaire were overlapping, the estimates from the questionnaire have been used to maintain consistency with the majority of other non-overlapping estimates, which come from the questionnaires. The actual variables in the model used the average of minutes per day the students spent on the material.

The diary was used to supplement and validate the questionnaire. The Spearman correlation coefficients between measures of practice estimated by questionnaires and diaries in the weeks when both were given (6th week of the semester) were 0.88 for DP and 0.73 for NDP in the Introduction to Psychology course. The overlap between the two instruments as measured in percentages (estimates diary / estimates questionnaire × 100) was 82% (median of all individual overlaps) for DP and 86% for NDP in the Introduction to Psychology course. The Statistics course (based on overlapping in the 6th and 13th week of the semester) had overlapping Spearman correlation coefficients of 0.75 for DP and 0.76 for NDP. The overlap measured by percentage had similar rates of 78% for DP and 88% for NDP. This high overlap is in line with previous results (Bilalić et al., 2007) and is evidence that the estimates about the previous week's activities from the questionnaire are valid.

Motivated Strategies for Learning Questionnaire – MSLQ

One can assume that the activity, which involves solitary learning without distraction (our definition of DP) is typical of self-regulated learning (Nandagopal & Ericsson, 2012a, b; Plant et al., 2005). We supplemented this measure of DP by the Motivated Strategies for Learning Questionnaire – MSLQ (Pintrich, 1991). Beside a motivational component (e.g., value beliefs, expectancy, and affect), the MSLQ also measures students' learning strategies for a particular course. The learning strategies captured by the MSLQ include cognitive strategies (e.g., rehearsal, elaboration, organization, and critical thinking), metacognition (e.g., ability to monitor their own mental processes and adjust them when needed), and resource management (e.g., time and study environment regulation, effort regulation, peer learning, and help seeking). We were particularly interested in metacognition, time and study environment regulation, and effort regulation, since they are the key components of self-regulated learning (Nandagopal & Ericsson, 2012a; Zimmerman, 2008). As such, they should differentiate between DP and NDP – they should be positively associated with DP but not associated (or even negatively associated) with NDP. The reliability of the individual MSLQ scales for Introduction to Psychology ranged between 0.55 and 0.80; and for Statistics between 0.62 and 0.82 (see SM, Sect. 3, for individual reliability estimates).

Previous knowledge tests

The knowledge test for the Introduction to Psychology course consisted of 35 (multiple choice) questions about general knowledge on psychology which is taught in high school psychology courses (compulsory for all high schools). The questions were also used for the entrance exams in previous generations and had reliability estimates (α) of 0.84. The knowledge test for the Statistics course contained 56 mathematical questions, starting with simple arithmetic exercises and ending with queries about percentages and probabilities. The items were chosen from the introductory chapters of the statistical textbooks (Petz et al., 2012), which covered the mathematical knowledge necessary for the introductory course on statistics in psychology. Mathematic is a compulsory throughout high school, but probability calculus (i.e. combinatorics) is not compulsory in most high schools. The reliability estimate (α) for the knowledge test in statistics was 0.89. In both knowledge tests, correctly answered questions were added and the final score reflected the number of correctly answered items.

Exams for Introduction to Psychology and Statistics

The content and form of the two individual exams for Introduction to Psychology and three exams for Statistics can be found in SM (see Table 1 for schedule). The individual exams in both Introduction to Psychology and Statistics were part of the Psychology curriculum which were given to every student generation and generally had a reliability of $\alpha > 0.75$ (for details about the reliability of the individual exams and their different parts, see SM, Sect. 3).

Difficulty of Introduction to Psychology and Statistics courses

Students answered three questions about the difficulty and complexity of the topics covered in the two courses repeatedly over the course of the semester. The first question asked how difficult they found the lectures using a scale from 1 to 7, where 1 is very easy and 7 very difficult. The second question was about the effort invested in understanding the topics (on a scale from 1 to 7 where 1 is no effort required, and 7 a lot of effort needed), while the final third question asked about satisfaction with the progress made for preparing for exams (using a scale from 1 to 5 where 1 is not satisfied at all and 5 is very satisfied). The questions were incorporated in the administered questionnaires and diaries.

Data Analysis

We first confirmed that Intelligence and WM are empirically better used separately then as a single construct using Confirmatory Factor Analysis (CFA – see SM, Sect. 3). We then created latent constructs of Performance (exams), Intelligence (intelligence tests), and Working Memory (WM) by using a Factor Analysis (FA) procedure. The FA confirmed a single factor structure for all constructs. We then used the factor loadings for individual variables in a regression model

to produce latent (and standardized) factors of Academic Performance, Intelligence, and Working Memory. The idea behind creating latent measures of constructs out of a few measurements of the same constructs is twofold: on the one hand it improves the capturing of the measured concepts (Conway et al., 2005; Đokić et al., 2018; Foster et al., 2015), and on the other, it reduces the complexity of the statistical model (Kline, 2015). The detailed factor loadings and percentage the new latent variables explained can be found in SM (Sect. 3).

All measures were normally distributed except the DP and NDP measures which were positively skewed. The Bayesian analysis used here is more robust to issues with non-normality (Kruschke, 2011), while the additional analysis in the frequentist mould used a bootstrapping procedure, a standard remedy for non-normally distributed measures (Pek et al., 2018), for the estimation of the standard errors of the variables.

For the main analysis, we decided to use conditional process analysis (Hayes, 2017; Shipley, 2016) in the Bayesian framework (Kruschke, 2011), which enabled us to test the theoretical m-FAP model. Figure 1 illustrates, for example, that the Intellectual Capacities were supposed to influence Knowledge and DP, which in turn influenced the AP. The conditional process analysis on a model set up this way allows for testing the indirect effect of Intellectual Capacities on AP through Knowledge and DP. In other words, one of the main theoretical questions, namely whether Knowledge and DP mediate the influence of Intellectual Capacities, can be answered with our model. Similarly, we could incorporate the influence of the different course in the path analysis by checking the interaction of the course factor with individual path relations – also known as moderated mediation (Hayes, 2017).

Bayesian analysis information

The Bayesian framework was chosen for its flexibility which enabled us to conduct all analyses within a single framework, as well as its ability to provide rich information about the model and its parameters (Kruschke, 2011). Here we provide a brief overview of the main options in the Bayesian analysis. The full information on the main aspects of the analysis is available in the SM (Sect. 5), while we also provide the data and code for the analysis (https://osf.io/9sr53/?view_only=5a0d2d53eb674673884aa090d36a1aa3).

The m-FAP model was modeled using the Gaussian function as the outcome variables (Academic Performance, Knowledge, and Deliberate/NonDeliberate Practice) were normally distributed. We used three different priors. The first one was the default (flat) prior, which was provided by the brms software (Bürkner, 2017). The second prior was an un-informative prior for all coefficients that followed the normal distribution with mean centered around 0 and standard deviation 1. This prior covered the whole range of plausible possibilities, with those between 1 and -1 being most probable (i.e. 68% of all possibilities), as all coefficients in our model were standardized (i.e. z-scale between -3 and 3). The final priors used the estimates from the studies mentioned in the power analysis section. Instead of using the exact estimates (e.g. mean and SD) from the previous studies, which would bias the results (Dienes, 2021), we used half the estimated coefficients from the previous studies, with the SD defined such that $M + 1SD$ reached the full estimate (i.e. if the coefficient was $+0.50$, the prior was $M = 0.25$ with $SD = 0.25$). We provide the results of the model with the informative priors in the main text. The results of the models with the un-informative and default priors, as well as other details about the priors, were presented in the SM within the sensitivity analysis (Sect. 5).

The estimated coefficients (e.g., reported central tendencies) were supplemented by four different indicators of existence and significance (Makowski et al., 2019). We first provide the credible intervals (CrI) of the posterior distribution. Here we use the common 95% CrI, as well as 89% CrI as suggested by some researchers (Kruschke, 2011; McElreath, 2018) and used in previous research (Bilalić et al., 2021; Vaci et al., 2019a) The coefficients that do not encompass zero within the CrI are highlighted in the figures (with "*" and "†" for the 95% and 89%CrI, respectively). The other measure of effect existence, probability of direction (pd), is also provided. The pd measure indicates how much the effect/coefficient is greater (or smaller if negative) than zero. More precisely, it presents the proportion of the posterior distribution that is of the same sign as the mean estimate and is a Bayesian equaling of the p measure in the frequentist statistics (Makowski et al., 2019).

Besides the two measures of consistency of an effect in one particular direction (CrI and pd), we present two measures of significance, that is, how important the estimates are: Regions Of Practical Equivalence (ROPE) and Bayes Factor (BF). The ROPE indicates that the values within the CrI are important by defining the region of unimportance and testing how much of the posterior distribution falls within that region of practical equivalence (Kruschke, 2018). Here we defined the minimal effect as a range which encompassed half of the nominally small effect (0.20) as recommended (Kruschke, 2018; p. 277). In other words, the range is -0.05 and 0.05. BF is the degree by which the probability mass has shifted away from or toward the null value, after observing the data.

Unlike the measures of existence, ROPE and BF can tell us not only whether the effect exists, but also whether the effect does not exist. The percentage of ROPE would be very high (e.g., 95%) and BF can be expressed in two measures, one for the strength of the support for the effect (BF_{10}) and the other for the strength of the support for the null

hypothesis (BF_{01}). The common interpretation of the BF is that the values between 1 and 3 are considered "anecdotical evidence", between 3 and 10 as "substantial", 10 – 30 "strong", 30 – 100 "very strong", and above 100 as "decisive" (Jarosz & Wiley, 2014).

Additional Analyses

We decided against using Structural Equation Modelling (SEM) as it is difficult to accommodate different number of exams (and consequently DP and NDP estimates) into a single model which would include both groups (there are two exams for Introduction to Psychology and three for Statistics). We do provide, however, separate SEM models for each of the two courses (SM, Sect. 8). We also ran the same model for the individual exams within course (SM, Sect. 9). These supplementary analyses confirm the pattern of results for the whole course. Finally, analyses in the frequentist framework were used to demonstrate that the same conclusions could be drawn using a different statistical approach (SM, Sect. 7).

Results

Descriptive Analysis

Course Difficulty Students found the Statistics course more difficult and challenging than the Introduction to Psychology: they judged the lectures more difficult ($M = 4.1$, $SD = 1.1$ vs. $M = 3.9$, $SD = 1.1$; $t(88) = 3.4$, $p = 0.01$, $d = 0.36$), they stated that they needed to invest more effort in understanding topics ($M = 4.5$, $SD = 1.2$ vs. $M = 4.3$,

$SD = 1.2$; $t(88) = 4.2$, $p < 0.001$, $d = 0.44$), and they were less satisfied with the progress they made in preparing for exams ($M = 3.5$, $SD = 0.6$ vs. $M = 3.6$, $SD = 0.6$; $t(105) = 4.4$, $p < 0.001$, $d = 0.43$).

Self-Regulated Learning and Practice The validity of the DP, here defined as solitary learning without distraction, is corroborated by its associations with the self-regulated learning strategies from the MSLQ. The key components of self-regulated learning such as metacognition, effort regulation, positive learning strategies, as well as time and study environment organization (Nandagopal & Ericsson, 2012a; Zimmerman, 2008) were all positively correlated with the DP measures in both courses (between $r = 0.21$ and 0.35). In contrast, no such connections were found with the NDP in both courses, as most of the self-regulated learning strategies were even negatively associated. Most importantly, the difference between correlations for DP and NDP were significant for key components of the self-regulated learning (see SM, Sect. 3).

Constructs Intercorrelations The descriptive statistics of the main variables presented in Table 2 show that both Intelligence and WM did not significantly correlate with the performance on the less challenging course, Introduction to Psychology. In the more challenging course, Statistics, both Intelligence and WM were significantly correlated with performance, with Intelligence being particularly highly correlated. Although previous knowledge was significantly related to performance in both courses, it followed a similar pattern as the intellectual capacity measures where it was a weaker predictor of the less challenging course than the more challenging one.

Table 2 Intercorrelations for Introduction to Psychology and Statistics courses

		Per	Intell	WM	Know	DP	NDP	Mean	SD
Performance	Intro							0.01	1.01
	Stats							-0.06	1.03
Intelligence	Intro	0.13						0.01	1.00
	Stats	**0.51***						0.01	1.00
WM	Intro	0.14	**0.48***					-0.02	1.01
	Stats	**0.39***	**0.48***					-0.02	1.01
Knowledge	Intro	**0.32***	**0.58***	**0.28***				16.95	4.34
	Stats	**0.62***	**0.57***	**0.37***				26.89	9.09
DP	Intro	**0.24***	-0.18	**-0.29***	-0.14			40.76	30.95
	Stats	-0.04	**-0.24***	**-0.23***	**-0.21***			22.48	22.57
NDP	Intro	0.04	0.16	0.17	0.16	**-0.45***		17.26	19.87
	Stats	0.16	0.07	0.09	0.08	-0.16		16.06	13.45

Note. Performance = z-average of exams, Intro = Introduction to Psychology, Stats = Statistics, Intelligence = Intelligence z-score average, WM = Working Memory z-score average, Knowledge = Score on the test of previous psychological or statistical knowledge, DP = Deliberate Practice (average minutes per day), NDP = Non-Deliberate Practice (average minutes per day). * $p < 0.05$.

In contrast, the amount of DP was significant for the less challenging course, whereas it was not for the more challenging one. It is important to note that both Intelligence and WM were significantly and positively correlated with previous knowledge in both courses. The same measures were negatively correlated with DP. DP and NDP are negatively associated, especially in the Introduction to Psychology course – students who engage in self-regulated learning activities unsurprisingly have less time for other, less useful learning activities.

Mediated–Factors of Academic Performance (m-FAP)

We used theoretical considerations (Fig. 1) when designing conditional process analysis for testing the interplay between different factors of academic performance. The model included direct connections between Intelligence, WM, previous knowledge, DP, and NDP towards the performance on a particular course. The indirect connections of Intelligence and WM to academic performance were also specified via previous knowledge and DP/NDP. Similarly, previous knowledge was also supposed to indirectly affect

performance via practice (both DP and NDP). Finally, the two courses (Introduction to Psychology and Statistics) were acting as moderators on all the specified relations, including mediations (i.e. moderated mediation, Hayes, 2017).

The model results summarized in Fig. 2, confirm the theoretical assumptions that Intellectual Capacities are influencing Academic Performance (AP) not only directly, but also indirectly through both Knowledge and Practice (for model fit indices, see SM, Sect. 5). There were, however, marked differences between the two courses (Fig. 2, path coefficients in blue for Introduction to Psychology and in red for Statistics). Intelligence was not a reliable direct predictor of AP in the Introduction to Psychology course (relation #1 in Table 3: $\beta = 0.02$; 95%CrI = -0.15 – 0.19; pd = 0.60; ROPE% = 42; $BF_{01} = 2.2$; $BF_{10} = 0.5$) but it was for Statistics ($\beta = 0.26$; 95%CrI = 0.08 – 0.44; pd = 1; ROPE% = 1; $BF_{01} = 0.1$; $BF_{10} = 16.2$), which produced a highly reliable pattern between courses in the Intelligence → AP relations (the difference between two coefficients are indicated by Δ in Fig. 2: $\beta = -0.24$; 95%CrI = -0.38 – -0.10; pd = 1; ROPE% = 0; $BF_{01} = 0$; $BF_{10} = 69.5$). Previous knowledge positively influenced AP in both courses, but it had a somewhat larger influence in Statistics (relation #3: $\beta = 0.34$; 95%CrI = 0.21 – 0.48; pd = 1; ROPE% = 0; $BF_{01} = 0$;

Fig. 2 Result Summary of the mediated–Factors of Academic Performance Model (m–FAP). The interplay between intellectual capacities (Intelligence and Working Memory), practice, Previous Knowledge, and Academic Performance for two academic courses: Introduction to Psychology and Statistics
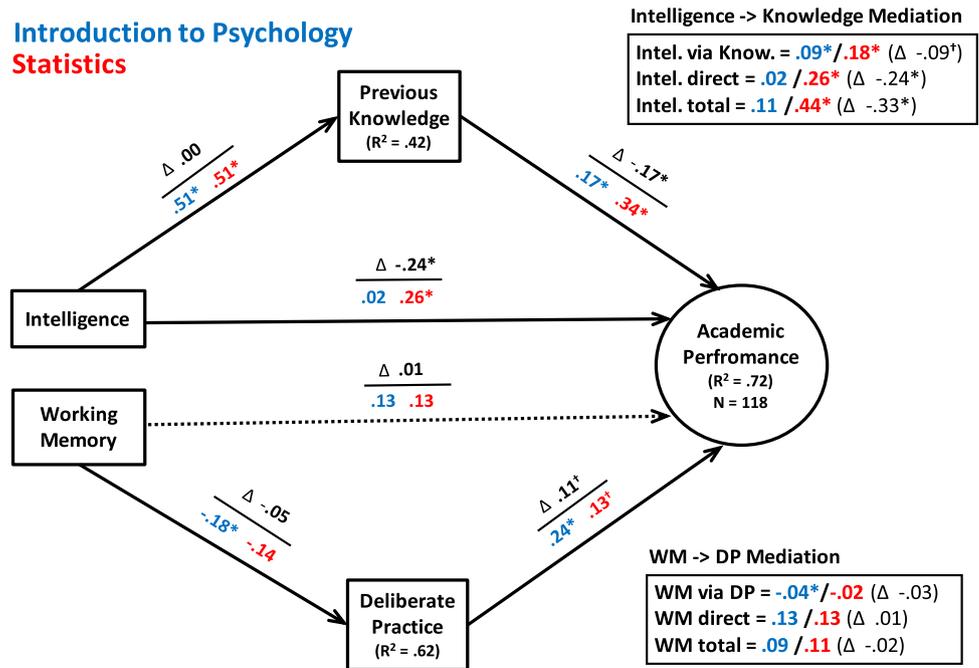
Note. The predictors are denoted by the rectangular shape while the dependent variable is denoted by the circular shape. Lines with single-end arrows indicate the direction of influence. The numbers on the line are standardized path model coefficients. The Introduction to Psychology path coefficients are in blue, the Statistics coefficients are in red, while their differences, indicated also by Δ, are in black. The indirect influence of Intelligence on Academic Performance through Previous Knowledge is formally tested in a mediation model (upper right box). The indirect influence of Working Memory on Academic Performance through Deliberate Practice is also formally tested by mediation (lower right box). R^2 is Bayesian full model coefficient of determination. *95% credible intervals do not encompass 0; †89% credible intervals do not encompass 0. Only statistically reliable relations were presented (see Table 3 below for the full model)

Table 3 Results of the m-FAP model (Fig. 2) for Introduction to Psychology and Statistics and direct comparisons of the main effects

# Relation		Indices of existence					Indices of significance		
A. Main Effects	Course	β	CrI low	-	high	pd	ROPE %	BF$_{01}$	BF$_{10}$
1 Intelligence → Performance	Intro	0.02	-0.15	-	0.19	0.60	42	2.2	0.5
	Stats	0.26*	0.08	-	0.44	1.00	1	0.1	16.2
	Diff	-0.24*	-0.38	-	-0.10	1.00	0	0.0	69.5
2 WM → Performance	Intro	0.13	-0.05	-	0.31	0.92	16	2.0	0.5
	Stats	0.13	-0.06	-	0.31	0.91	18	2.4	0.4
	Diff	0.01	-0.15	-	0.16	0.53	48	2.5	0.4
3 Knowledge → Performance	Intro	0.17*	0.05	-	0.30	1.00	3	0.1	8.9
	Stats	0.34*	0.21	-	0.48	1.00	0	0.0	> 1000
	Diff	-0.17*	-0.30	-	-0.03	0.99	4	0.1	7.8
4 DP → Performance	Intro	0.24*	0.11	-	0.37	1.00	0	0.0	118.7
	Stats	0.13†	-0.01	-	0.27	0.97	11	0.4	2.3
	Diff	0.11	-0.02	-	0.24	0.96	17	0.5	2.1
5 NDP → Performance	Intro	0.05	-0.09	-	0.20	0.77	41	5.4	0.2
	Stats	0.07	-0.06	-	0.21	0.86	34	4.4	0.2
	Diff	-0.02	-0.17	-	0.13	0.60	47	2.5	0.4
6 Intelligence → Knowledge	Intro	0.51*	0.35	-	0.67	1.00	0	0.0	> 1000
	Stats	0.51*	0.34	-	0.68	1.00	0	0.0	> 1000
	Diff	0.00	-0.19	-	0.19	0.50	39	2.0	0.5
7 WM → Knowledge	Intro	0.03	-0.16	-	0.20	0.62	39	5.1	0.2
	Stats	0.12	-0.07	-	0.30	0.89	20	2.6	0.4
	Diff	-0.09	-0.30	-	0.11	0.81	26	1.3	0.8
8 Intelligence → DP	Intro	-0.07	-0.25	-	0.11	0.77	33	1.7	0.6
	Stats	-0.11	-0.30	-	0.07	0.88	21	1.1	0.9
	Diff	0.05	-0.09	-	0.18	0.75	43	1.2	0.9
9 WM → DP	Intro	-0.18*	-0.36	-	-0.01	0.98	6	0.2	4.1
	Stats	-0.14	-0.32	-	0.04	0.94	14	0.8	1.3
	Diff	-0.05	-0.18	-	0.09	0.75	44	1.2	0.9
10 Knowledge → DP	Intro	-0.03	-0.17	-	0.11	0.65	48	2.6	0.4
	Stats	-0.04	-0.20	-	0.11	0.72	42	2.5	0.4
	Diff	0.02	-0.13	-	0.16	0.59	49	1.3	0.8
11 Intelligence → NDP	Intro	0.06	-0.13	-	0.24	0.73	34	1.7	0.6
	Stats	0.03	-0.16	-	0.23	0.62	37	2.1	0.5
	Diff	0.03	-0.12	-	0.17	0.64	48	1.3	0.8
12 WM → NDP	Intro	0.10	-0.08	-	0.27	0.86	25	1.2	0.8
	Stats	0.08	-0.11	-	0.27	0.80	29	1.7	0.6
	Diff	0.02	-0.12	-	0.16	0.60	50	1.4	0.7
13 Knowledge → NDP	Intro	0.04	-0.12	-	0.20	0.68	42	2.2	0.5
	Stats	0.00	-0.16	-	0.16	0.51	46	2.7	0.4
	Diff	0.04	-0.11	-	0.19	0.68	44	1.2	0.9

Note. Intro = Introduction to Psychology, Stats = Statistics, Intelligence = Intelligence z-score average, WM = Working Memory z-score average, Knowledge = Score on the test of previous knowledge, DP = Deliberate Practice average minutes per day, NDP = Non-Deliberate Practice average minutes per day, Performance = z-average of exams, → = direction of influence, Std. B = standardized path coefficients, CrI = 95% Credible Intervals, pd = Proportion of Direction, ROPE% = % in the Region of Practical Equalence (-0.05 to +0.05), BF = Bayes Factor. Terms in the squared parenthesis in sections B and C specify how the particular coefficients were obtained. *95% Credible Intervals (CrI) do not encompass 0, †89% Credible Intervals (CrI) do not encompass 0.

BF$_{10}$ > 1000) than in Introduction to Psychology (β = 0.17; 95%CrI = 0.05 – 0.30; pd = 1; ROPE% = 3; BF$_{01}$ = 0.1; BF$_{10}$ = 8.9). In contrast, DP displayed a different pattern of

results: DP was a reliable and significant direct predictor of AP for Introduction to Psychology (relation #4: β = 0.24; 95%CrI = 0.11 – 0.37; pd = 1; ROPE% = 0; BF$_{01}$ = 0;

$BF_{10} = 118.7$), but only borderline for Statistics ($\beta = 0.13$; 95%CrI = -0.001 – 0.27; pd = 0.97; ROPE% = 11; $BF_{01} = 0.4$; $BF_{10} = 2.3$).

Both intellectual capacities, Intelligence and WM, also indirectly influenced AP through two positively significant predictors, Previous Knowledge and DP, respectively (Fig. 2). More intelligent students tended to have more previous knowledge, which then positively influenced their academic performance. This mediation pattern of Intelligence's indirect influence on AP through Knowledge was formally tested and found statistically reliable (Fig. 2, upper right box and relation #14 in Table 4) for Statistics ($\beta = 0.18$; 95%CrI = 0.09 – 0.27; pd = 1; ROPE% = 0; $BF_{01} = 0$; $BF_{10} > 1000$) and to an extent also in Introduction to Psychology ($\beta = 0.09$; 95%CrI = 0.02 – 0.17; pd = 1; ROPE% = 14; $BF_{01} = 0$; $BF_{10} = 34$). The somewhat smaller, but reliable and significant, mediation effect in Introduction to Psychology compared to Statistics ($\beta = 0.09$; 95%CrI = -0.18 – 0.001; pd = 0.97; ROPE% = 22; $BF_{01} = 0.2$; $BF_{10} = 5.6$) was a consequence of smaller influence of Knowledge on AP in Introduction to Psychology. Intelligence had a large positive effect on Previous Knowledge in

both courses (0.55 and 0.55 for Introduction to Psychology and Statistics, respectively).

WM yielded a different pattern because the students with more WM capacity were inevitably investing less in DP. This in turn reduced the overall direct influence of WM on academic performance. This mediation was, however, only reliable in Introduction to Psychology (relation #17 in Table 4: $\beta = -0.04$; 95%CrI = -0.10 – -0.001; pd = 0.98; ROPE% = 62; $BF_{01} = 0.1$; $BF_{10} = 10.5$) as the relations in the indirect path, WM → DP → AP, were only reliable for Introduction to Psychology and not Statistics (see Fig. 1 and its lower right box).

The total overall effects of Intelligence on AP, which also includes the mediations (through Previous Knowledge), were large and statistically reliable for Statistics (relation #16 in Table 4: $\beta = 0.44$; 95%CrI = 0.25 – 0.62; pd = 1; ROPE% = 0; $BF_{01} = 0$; $BF_{10} > 1000$) unlike for Introduction to Psychology ($\beta = 0.11$; 95%CrI = -0.05 – 0.27; pd = 1; ROPE% = 21; $BF_{01} = 1.1$; $BF_{10} = 1$), resulting in a consistent difference between the two courses (see the right upper box in Fig. 2: $\beta = -0.33$; 95%CrI = -0.47 – -0.18; pd = 1; ROPE% = 0; $BF_{01} = 0$; $BF_{10} = 982$). The total overall effects of WM on

Table 4 Results of the m-FAP model (Fig. 2) for Introduction to Psychology and Statistics and direct comparisons of the mediation/indirect effects

# Relation			Indices of existence				Indices of significance		
B. Mediation: Intelligence via *Knowledge*	*Course*	β	CrI low	-	high	pd	ROPE %	BF_{01}	BF_{10}
14 Indirect/mediated Intel. [6×3]	Intro	0.09*	0.02	-	0.17	1.00	14	0.0	34.7
	Stats	0.18*	0.09	-	0.27	1.00	0	0.0	> 1000
	Diff	-0.09†	-0.18	-	0.00	0.97	22	0.2	5.6
15 Direct Intelligence [1]	Intro	0.02	-0.15	-	0.19	0.00	42	0.0	0.0
	Stats	0.26*	0.08	-	0.44	0.00	1	0.0	0.0
	Diff	-0.24*	-0.38	-	0.10	0.00	0	0.0	0.0
16 Total Intelligence [1+(6×3)]	Intro	0.11	-0.05	-	0.27	1.00	21	1.1	1.0
	Stats	0.44*	0.25	-	0.62	0.91	0	0.0	> 1000
	Diff	-0.33*	-0.47	-	0.18	1.00	0	0.0	982.2
C. Mediation: WM via *DP*									
17 Indirect/mediated WM [9×4]	Intro	-0.04*	-0.10	-	0.00	0.98	62	0.1	10.5
	Stats	-0.02	-0.06	-	0.01	0.91	95	0.7	1.4
	Diff	-0.03	-0.07	-	0.01	0.93	88	0.4	2.3
18 Direct WM [2]	Intro	0.13	-0.05	-	0.31	0.00	16	2.03	0.49
	Stats	0.13	-0.06	-	0.31	0.00	18	2.41	0.42
	Diff	0.01	-0.15	-	0.16	0.00	48	2.54	0.39
19 Total WM [2+(9×4)]	Intro	0.09	-0.10	-	0.27	0.87	28	3.6	0.3
	Stats	0.11	-0.08	-	0.29	0.82	23	3.0	0.3
	Diff	-0.02	-0.18	-	0.13	0.60	46	2.5	0.4

Note. Intro = Introduction to Psychology, Stats = Statistics, Intelligence = Intelligence z-score average, WM = Working Memory z-score average, Knowledge = Score on the test of previous knowledge, DP = Deliberate Practice average minutes per day, NDP = Non-Deliberate Practice average minutes per day, Performance = z-average of exams, → = direction of influence, Std. B = standardized path coefficients, CrI = 95% Credible Intervals, pd = Proportion of Direction, ROPE% = % in the Region of Practical Equalence (-0.05 to +0.05), BF = Bayes Factor. Terms in the squared parenthesis in sections B and C specify how the particular coefficients were obtained. *95% Credible Intervals (CrI) do not encompass 0, †89% Credible Intervals (CrI) do not encompass 0.

AP (including mediations DP) were not statistically reliable in either of the courses (see the upper right box in Fig. 2).

Figure 2 presents only the statistically reliable relations in the m-FAP model. Tables 3 and 4 present the coefficients for all relations in both courses, as well as their differences. For the sake of simplicity, Table 3 presents the information for the main effect only (arrows in Fig. 2), while Table 4 reports the indirect/mediation and total effects (boxes in Fig. 2). For example, the assumed relation between Intelligence and Academic Performance (Relation #1 in Table 3, Intelligence→Performance), is given for Introduction (Std. B = 0.02), Statistics (Std. B = 0.26), and their difference (Std. B = -0.24). Both credible intervals and pd indicate the existence of the effect for Statistics and the differences, while ROPE and BF_{10} confirm that these effects are also significant. On the other hand, the Intelligence→Performance relation for the Introduction to Psychology course is certainly not significant, but it is not possible to claim that it does not exists (both ROPE and BF_{01} indicators are ambiguous).

Of theoretical interests would be the fact that the assumed relation between Previous Knowledge and Practice (see relation #10 in Table 3), where more knowledgeable students tend to practice less, was not reliable in either of the courses. Similarly, it is notable that NDP is ineffective as a predictor of academic performance and has no other significant influences (relation #5). Based on the Bayes Factor (see BF_{01} in Table 3), however, it is not easy to conclude that Knowledge→DP is not significant as the BF_{01} is below the threshold for weak evidence (i.e. 3). The NDP→Performance relation does indeed look to be non-existent given that the size of the BF_{01} for both courses (BF_{01} = 5.4 and BF_{01} = 4.4, for Introduction to Psychology and Statistics, respectively).

Discussion

The results confirm previous theoretical assumptions of a complex interplay between the factors contributing to academic performance. They also extend the theory behind the factors of academic performance, as they specify how particular needs change the dynamics between individual factors on academic performance. Intellectual capacities were related to academic performance directly, but also indirectly through previous knowledge and practice. The extent of the direct and indirect relations depended on the context. In a novel, more challenging course, where logical thinking and understanding of complex concepts was necessary, intellectual capacities, in particular Intelligence, still exerted a direct influence on academic performance. In a familiar, less challenging environment, where accumulation of knowledge and its reproduction in exams was required, its influence was more of an indirect nature. Here, other factors, such as effort and goal-oriented learning (DP), were more important. The same factor of practice was, however, less predictive of academic performance in a more challenging setting.

Intellectual capacities

Our study demonstrated why it is difficult to pinpoint the influence of intellectual capacities on academic performance. In a course where accumulation of knowledge and its reproduction in exams is necessary, intellectual capacities did not directly determine performance. The process of acquiring facts and reproducing them by rote, as was the case in the exams in the Introduction to Psychology course, requires more focused effort and practice than it does intellectual capacities. However, Intelligence still indirectly influenced performance in the Introduction to Psychology course through previous knowledge (at least reliably at the 89%CrI – see Table 3). More intelligent students tend to be more intellectually curious and are more likely to acquire knowledge about their future subjects even before the actual courses begin. The acquired knowledge then positively predicted their performance in the course, but it also probably weakened the direct influence of intelligence (similar as to the role of DP in the WM – academic performance relation).

The other course, Statistics, required different kinds of abilities, as one could not rely on rote memorization. Here, understanding concepts and the connections between them was a prerequisite for success. It is not surprising, then, that Intelligence, which captured abstract reasoning and logic, was a considerably better predictor than WM. Intelligence influenced performance in Statistics directly even after it was mediated through previous knowledge. As in the other course, Introduction to Psychology, previous knowledge was highly and positively predictive of performance. Again, as in the other course, it was significantly determined by Intelligence – more intelligent students had acquired more relevant knowledge.

Unlike with the Introduction to Psychology course, WM was neither directly nor indirectly related to performance in the Statistics course. One reason is certainly that WM correlates highly with Intelligence (see Table 2), which significantly influences performance. Once one construct explains the shared variance, as Intelligence does here, the other related concept, in this case WM, will inevitably have less variance to explain despite being correlated with the construct (see Table 2). The structure of the exams could also provide an explanation for different patterns of relations that Intelligence and WM have with academic performance in the Introduction to Psychology and Statistics. Introduction to Psychology exams consisted not only of factual multiple-choice questions, but also of true/false, open-ended,

gap-filling, and matching questions. These types of test questions allowed for WM to exert its functions, at least marginally above and beyond (over)learning the facts due to extensive practice throughout the semester. For example, WM could keep relevant information active in primary memory while sketching answers to the open-ended questions or managing strategic search through the long-term memory in the presence of intrusions while solving true/false or gap-filling exercises (Unsworth, 2016). In contrast, Statistics exams put more burden on the problem solving and if–then type of reasoning (both in theoretical and framework of solving applied statistical problems) than on the mere recollection of previously learned facts.

Exam structure of this kind emphasized generating and testing alternative hypotheses which could provide space for Intelligence to be not only indirectly, but also directly related to performance (Engle, 2018; Shipstead et al., 2016). In addition, in the Statistics exams students had most of the relevant information presented externally (e.g. all necessary statistical formulas for statistical exercises as well as alternative answers in multiple-choice problems were provided in the exam materials). This furthermore reduced their WM load and consequently rendered its influence nugatory. As much as this interpretation is speculative and requires further empirical testing, inclusion of separate measures of Intelligence and WM in the study provided us with a welcome opportunity to explore the possibility of different roles that these two constructs might have in complex cognition (Burgoyne et al., 2019a; Engle, 2018; Shipstead et al., 2016).

Previous (domain-specific) knowledge

Previous knowledge was predictive of success in both courses. The more students knew about psychology and statistics, the better they performed. This is expected, as it is easier to understand, remember, and recall incoming information if there is an already existing knowledge base which can accommodate the incoming material (Atkinson & Shiffrin, 1971; Shiffrin & Atkinson, 1969). It is plausible that students are less likely to come to the courses with an already solid understanding of mathematical and statistical concepts (e.g., probability) than with factual knowledge about psychology in general. This would mean that practice itself could not completely compensate for the lack of previous knowledge in statistics. The short duration of the course, just a single semester, would also make it difficult for students who came with insufficient mathematical knowledge to compensate for this through hard work. This may also have practical implications as the statistics curriculum in the study of psychology may benefit from more attention devoted to basic concepts of probability before moving to other more specific concepts.

The interplay between learning context (Deslauriers et al., 2019) and knowledge may go a long way towards explaining previous contradictory results when it comes to the role of knowledge in academic performance (Masui et al., 2014; Nandagopal & Ericsson, 2012b). Most previous studies took the average grade in complete studies as their performance measure (Nonis & Hudson, 2006; Richardson et al., 2012; Vedel, 2014). Similarly, the previous knowledge was often just a grade average from high school (Diseth, 2011; Richardson et al., 2012). Our study shows that the role of knowledge is probably underestimated in those studies. Once knowledge was kept specific to a particular course, and once this knowledge was clearly specified with a tailor-made pretest, previous knowledge was arguably the most important predictor of success.

(Deliberate) Practice

Practice was an important factor in academic performance in both courses. However, only activities that avoided learning with distractions, whether external ones such as TV and the internet, or the presence of other people, were efficient. In other words, self-regulated learning of the material was the only positive predictor of academic success. This has been confirmed here, as the learning strategies, which include specific behavior (e.g. elaboration, rehearsal) as well as monitoring one's own progress (e.g. metacognitive self-regulation) and resource management (e.g. environment and time of study, effort regulation), were positively associated with DP but not with NDP. The removal of distractors should self-evidently be beneficial for effectiveness in learning. It also indicates that these students are aware of challenges and consequently plan their learning better. Previous studies demonstrate that these metacognitive strategies of students who engage in self-regulated practice are well developed (Bol et al., 2016; Joseph, 2009). These students can judge what kind of material they need for full understanding, unlike their less metacognitively able colleagues. The result is engagement in this highly structured kind of practice, akin to DP in expertise domains (Nandagopal & Ericsson, 2012a, b).

The differentiation between focused learning activities and other, less efficient ones has been shown in previous studies (Nandagopal & Ericsson, 2012a; Plant et al., 2005). Our study goes beyond such studies as it demonstrates the situations where focused practice is more or less efficient. Practice was predictive of success in Introduction to Psychology; in Statistics, however, it was not a relevant predictor. The pattern is exactly opposite to that of intellectual capacities, which indicated the complementary influence and the interplay between practice and intellectual capacities. This could indicate that the nature of the Introduction to Psychology course, where one needed to acquire large amounts

of factual knowledge, allows students who are focusing their efforts to be rather efficient. In contrast, Statistics, where deeper understanding of concepts and their relations is necessary, makes it possible that even focused practice does not produce immediate result. This is evident in the relation of previous knowledge to performance, which was higher in Statistics than in the Introduction to Psychology course. It is plausible that the sheer difference in previous knowledge, coupled with the short duration of the course, made practice less important. Less knowledgeable students were simply too far behind initially to be able to catch up to their more knowledgeable peers in such a short time frame, despite the greater effort they invested.

Limitations

It should be noted that the learning activities that we label DP do not necessarily encompass immediate feedback from a teacher, which is considered a crucial aspect of DP (Debatin et al., 2021; Ericsson & Harwell, 2019; Ericsson et al., 1993; Ericsson, 2020, 2021). It is reasonable to assume that the learning activities with immediate feedback provided by an instructor would further strengthen the impact of practice on academic performance. However, such learning episodes are rare in the academic setting, as is seen in our study where only a handful of students used tutorials, i.e., one on one sessions with an instructor. Future studies may provide direct interventions that include individual sessions with instructors as a way of capturing these rare occasions.

Intellectual capacities, previous knowledge, and (deliberate) practice explained a respectable amount of variance in academic performance – almost three quarters (see Fig. 2). The unexplained variance, however, indicates that there may be factors that play important roles in academic achievement but that were not included in the study. For example, motivation (Linnenbrink-Garcia et al., 2018), academic-self efficiency (Honicke & Broadbent, 2016) or personality factors such as self-control (Duckworth et al., 2012, 2016), grit (Dixson, 2021; Duckworth et al., 2007; Rimfeld et al., 2016), or growth mind-set (Wang et al., 2022; Yeager et al., 2016) may additionally explain academic performance. Their influence on academic performance may be direct, where they explain additional variance on top of other factors (e.g., intelligence, see Duckworth et al., 2019). A more realistic path for their influence would be via other factors, such as practice. For example, grittier students may tend to engage more in (deliberate) practice, which in turn would lead to better results (Duckworth et al., 2011; Lee & Sohn, 2017). Similarly, students with greater self-control may be able

to better tune out distractions during learning, while students with growth mind-set may be more persistent. In all instances, these factors are easily incorporated in the m-FAP model and tested using the same conditional process analysis we have employed in this study.

The relatively small sample size in the study is offset using Bayesian analysis (Kruschke, 2011; McElreath, 2018) and informative priors (Smid et al., 2020; Zonder-van-Zwijnenburg et al., 2017), which proved sensitive enough (see the sensitivity analysis, SM, Sect. 5.3). More problematic is the fact that our study has featured only a single cohort of predominantly female students which was tested on two academic courses. It is obviously unclear how our results would transfer in different settings with more balanced samples. Our study, however, goes beyond empirical contribution as it provides an analytic tool, m-FAP model, with a cutting-edge statistical framework (Bayesian conditional process analysis) which can accommodate a wide range of factors such as gender or different courses.

Conclusions

Our study underlined that it is necessary to capture relevant factors of academic performance as they form complex interrelations, which are highly dependent on the needs and requirements of the individual courses. In our particular case, it meant capturing the complementary effects of intelligence, working memory, previous knowledge, and (non) deliberate practice on academic performance (Duckworth et al., 2019; Hambrick et al., 2018; Moreau et al., 2019; Vaci et al., 2019a, b). In case where intelligence is necessary, (deliberate) practice was not enough, whereas in a setting where intelligence had less influence, (deliberate) practice filled the gap. Future studies should take into account other factors of academic performance in the hope of uncovering a more complete picture of the interplay between the predictors of academic performance.

Declarations

Conflicts of interest/Competing interests Authors declare that there are no competing interests.

Ethics approval The research protocol was approved by the internal ethics committee of the Department of Psychology at the University of Sarajevo.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131*(1), 30.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American, 225*(2), 82–90. https://doi.org/10.1038/scientificamerican0871-82

Barrouillet, P., & Lépine, R. (2005). Working memory and children's use of retrieval to solve addition problems. *Journal of Experimental Child Psychology, 91*(3), 183–204.

Bergman Nutley, S., & Söderqvist, S. (2017). How is working memory training likely to influence academic performance? Current evidence and methodological considerations. *Frontiers in Psychology, 8*, 69.

Bilalić, M. (2017). *The neuroscience of expertise*. Cambridge University Press.

Bilalić, M., Gula, B., & Vaci, N. (2021). Home advantage mediated (HAM) by referee bias and team performance during covid. *Scientific Reports, 11*(1), 1–13.

Bilalić, M., McLeod, P., & Gobet, F. (2007). Does chess need intelligence?—A study with young chess players. *Intelligence, 35*(5), 457–470.

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences, 29*(2), 109.

Bloom, B. S. (1976). *Human characteristics and school learning*. McGraw-Hill.

Bol, L., Campbell, K. D. Y., Perez, T., & Yen, C.-J. (2016). The effects of self-regulated learning training on community college students' metacognition and achievement in developmental math courses. *Community College Journal of Research and Practice, 40*(6), 480–495. https://doi.org/10.1080/10668926.2015.1068718

Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science, 236*(4798), 157–161.

Burgoyne, A. P., Hambrick, D. Z., & Altmann, E. M. (2019a). Is working memory capacity a causal factor in fluid intelligence? *Psychonomic Bulletin & Review, 26*(4), 1333–1339.

Burgoyne, A. P., Harris, L. J., & Hambrick, D. Z. (2019b). Predicting piano skill acquisition in beginners: The role of general intelligence, music aptitude, and mindset. *Intelligence, 76*, 101383. https://doi.org/10.1016/j.intell.2019.101383

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1–28.

Castejón, J. L., Gilar, R., & Pérez, A. M. (2006). Complex learning: The role of knowledge, intelligence, motivation and learning strategies. *Psicothema, 18*(4), 679–685.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*(1), 55–81.

Chinn, D., Sheard, J., Carbone, A., & Laakso, M.-J. (2010). Study habits of CS1 students: What do they do outside the classroom? *Proceedings of the Twelfth Australasian Conference on Computing Education-Volume, 103*, 53–62.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786.

Deary, I. J. (2011). Intelligence. *Annual Review of Psychology, 63*(1), 453–482. https://doi.org/10.1146/annurev-psych-120710-100353

Debatin, T., Harder, B., & Ziegler, A. (2019). Does fluid intelligence facilitate the learning of English as a foreign language?—A longitudinal latent growth curve analysis. *Learning and Individual Differences, 70*, 121–129.

Debatin, T., Hopp, M. D. S., Vialle, W., & Ziegler, A. (2021). The meta-analyses of deliberate practice underestimate the effect size because they neglect the core characteristic of individualization—An analysis and empirical evidence. *Current Psychology*. https://doi.org/10.1007/s12144-021-02326-x

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences, 116*(39), 19251–19257. https://doi.org/10.1073/pnas.1821936116

Dienes, Z. (2021). How to use and report Bayesian hypothesis tests. *Psychology of Consciousness: Theory, Research, and Practice, 8*(1), 9.

Diseth, Å. (2011). Self-efficacy, goal orientations and learning strategies as mediators between preceding and subsequent academic achievement. *Learning and Individual Differences, 21*(2), 191–195.

Diseth, Å., Pallesen, S., Brunborg, G. S., & Larsen, S. (2010). Academic achievement among first semester undergraduate psychology students: The role of course experience, effort, motives and learning strategies. *Higher Education, 59*(3), 335–352.

Dixson, D. D. (2021). Is grit worth the investment? How grit compares to other psychosocial factors in predicting achievement. *Current Psychology, 40*(7), 3166–3173. https://doi.org/10.1007/s12144-019-00246-5

Đokić, R., Koso-Drljević, M., & Đapo, N. (2018). Working memory span tasks: Group administration and omitting accuracy criterion do not change metric characteristics. *PLoS ONE, 13*(10), e0205169.

Dochy, F., De Rijdt, C., & Dyck, W. (2002). Cognitive prerequisites and learning: How far have we progressed since Bloom? Implications for educational practice and teaching. *Active Learning in Higher Education, 3*(3), 265–284.

Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research, 69*(2), 145–186.

Doumen, S., Broeckmans, J., & Masui, C. (2014). The role of self-study time in freshmen's achievement. *Educational Psychology, 34*(3), 385–402.

Duckworth, A. L., Kirby, T. A., Tsukayama, E., Berstein, H., & Ericsson, K. A. (2011). Deliberate practice spells success: Why grittier competitors triumph at the National Spelling Bee. *Social Psychological and Personality Science, 2*(2), 174–181.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087

Duckworth, A. L., Quinn, P. D., & Tsukayama, E. (2012). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology, 104*(2), 439–451.

Duckworth, A. L., Quirk, A., Gallop, R., Hoyle, R. H., Kelly, D. R., & Matthews, M. D. (2019). Cognitive and noncognitive predictors of success. *Proceedings of the National Academy of Sciences, 116*(47), 23499–23504. https://doi.org/10.1073/pnas.1910510116

Duckworth, A. L., White, R. E., Matteucci, A. J., Shearer, A., & Gross, J. J. (2016). A stitch in time: Strategic self-control in high school and college students. *Journal of Educational Psychology, 108*(3), 329–341.

Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science, 13*(2), 190–193.

Ericsson, K. A. (2020). Towards a science of the acquisition of expert performance in sports: Clarifying the differences between deliberate practice and other types of practice. *Journal of Sports Sciences, 38*(2), 159–176. https://doi.org/10.1080/02640414.2019.1688618

Ericsson, K. A. (2021). Given that the detailed original criteria for deliberate practice have not changed, could the understanding of this complex concept have improved over time? A response to Macnamara and Hambrick (2020). *Psychological Research, 85*(3), 1114–1120.

Ericsson, K. A., & Harwell, K. (2019). Deliberate practice and proposed limits on the effects of practice on the acquisition of expert performance: Why the original definition matters and recommendations for future. *Frontiers in Psychology, 10,* 2396.

Ericsson, K. A., Krampe, R. T., & Tesch-Roemer, C. (1993). The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review, 100,* 363–363.

Federici, L., & Schuerger, J. (1976). High school psychology students versus non-high school psychology students in a college introductory class. *Teaching of Psychology, 3*(4), 172–174.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition, 43*(2), 226–236.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science, 15*(6), 373–378.

Gagné, F., & St Père, F. (2001). When IQ is controlled, does motivation still predict achievement? *Intelligence, 30*(1), 71–100.

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A.-M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93*(3), 265–281.

Gobet, F., & Simon, H. A. (1996). Templates in Chess Memory: A Mechanism for Recalling Several Boards. *Cognitive Psychology, 31*(1), 1–40.

Griggs, R. A., & Jackson, S. L. (1988). A reexamination of the relationship of high school psychology and natural science courses to performance in a college introductory psychology class. *Teaching of Psychology, 15*(3), 142–144.

Hambrick, D. Z. (2003). Why are some people more knowledgeable than others? A longitudinal study of knowledge acquisition. *Memory & Cognition, 31*(6), 902–917.

Hambrick, D. Z., Burgoyne, A. P., Macnamara, B. N., & Ullén, F. (2018). Toward a multifactorial model of expertise: Beyond born versus made. *Annals of the New York Academy of Sciences, 1423*(1), 284–295.

Hambrick, D. Z., Pink, J. E., Meinz, E. J., Pettibone, J. C., & Oswald, F. L. (2008). The roles of ability, personality, and interests in acquiring current events knowledge: A longitudinal study. *Intelligence, 36*(3), 261–278.

Havighurst, R. J. (1953). *Human development and education*. Longmans, Green.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.

Heitz, R. P., Redick, T. S., Hambrick, D. Z., Kane, M. J., Conway, A. R., & Engle, R. W. (2006). Working memory, executive function, and general fluid intelligence are not the same. *Behavioral and Brain Sciences, 29*(2), 135.

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review, 17,* 63–84.

Horn, J. L., & McArdle, J. J. (2007). Understanding human intelligence since Spearman. *Factor Analysis at, 100,* 205–247.

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving, 7*(1), 2.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Praeger Publishers/Greenwood Publishing Group Inc.

Joseph, N. (2009). Metacognition Needed: Teaching Middle and High School Students to Develop Strategic Learning Skills. *Preventing School Failure: Alternative Education for Children and Youth, 54*(2), 99–103. https://doi.org/10.1080/10459880903217770

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Kruschke, J. K. (2011). *Doing Bayesian data analysis: A tutorial with R and BUGS*. Academic Press.

Kruschke, J. K. (2018). Rejecting or accepting parameter values in Bayesian estimation. *Advances in Methods and Practices in Psychological Science, 1*(2), 270–280.

Kyllonen, P. C. (2002). g: Knowledge, speed, strategies, or working-memory capacity? A systems perspective. In *The general factor of intelligence* (pp. 427–458). Psychology Press.

Lee, S., & Sohn, Y. W. (2017). Effects of grit on academic achievement and career-related attitudes of college students in Korea. *Social Behavior and Personality: An International Journal, 45*(10), 1629–1642.

Limanond, T., Jomnonkwao, S., Watthanaklang, D., Ratanavaraha, V., & Siridhara, S. (2011). How vehicle ownership affect time utilization on study, leisure, social activities, and academic performance of university students? A case study of engineering freshmen in a rural university in Thailand. *Transport Policy, 18*(5), 719–726.

Linnenbrink-Garcia, L., Wormington, S. V., Snyder, K. E., Riggsbee, J., Perez, T., Ben-Eliyahu, A., & Hill, N. E. (2018). Multiple pathways to success: An examination of integrative motivational profiles among upper elementary and college students. *Journal of Educational Psychology, 110*(7), 1026–1048.

Loftus, E. F., & Marburger, W. (1983). Since the eruption of Mt. St. Helens, has anyone beaten you up? Improving the accuracy of retrospective reports with landmarkevents. *Memory & Cognition, 11*(2), 114–120.

Macnamara, B. N., & Hambrick, D. Z. (2021). Toward a cumulative science ofexpertise: Commentary on Moxley, Ericsson, and Tuffiash (2017). *Psychological Research, 85*(3), 1108–1113.

Macnamara, B. N., & Maitra, M. (2019). The role of deliberate practice in expert performance: Revisiting Ericsson, Krampe

& Tesch-Römer (1993). *Royal Society Open Science, 6*(8), 190327.

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science, 25*(8), 1608–1618.

Macnamara, B. N., Moreau, D., & Hambrick, D. Z. (2016). The relationship between deliberate practice and performance in sports: A meta-analysis. *Perspectives on Psychological Science, 11*(3), 333–350.

Makowski, D., Ben-Shachar, M. S., Chen, S. A., & Lüdecke, D. (2019). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology, 10*, 2767.

Masten, A. S., & Coatsworth, J. D. (1998). The development of competence in favorable and unfavorable environments: Lessons from research on successful children. *American Psychologist, 53*(2), 205.

Masui, C., Broeckmans, J., Doumen, S., Groenen, A., & Molenberghs, G. (2014). Do diligent students perform better? Complex relations between student and course characteristics, study time, and academic performance in higher education. *Studies in Higher Education, 39*(4), 621–643.

McElreath, R. (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.

Moreau, D., Macnamara, B. N., & Hambrick, D. Z. (2019). Overstating the Role of Environmental Factors in Success: A Cautionary Note. *Current Directions in Psychological Science, 28*(1), 28–33. https://doi.org/10.1177/0963721418797300

Mosing, M. A., Hambrick, D. Z., & Ullén, F. (2019). Predicting musical aptitude and achievement: Practice, teaching, and intelligence. *Journal of Expertise, 2*(3), 184–197.

Nandagopal, K., & Ericsson, K. A. (2012a). An expert performance approach to the study of individual differences in self-regulated learning activities in upper-level college students. *Learning and Individual Differences, 22*(5), 597–609.

Nandagopal, K., & Ericsson, K. A. (2012b). Enhancing students' performance in traditional education: Implications from the expert performance approach and deliberate practice. In *APA educational psychology handbook, Vol 1: Theories, constructs, and critical issues* (pp. 257–293). American Psychological Association.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., & Sternberg, R. J. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77–101.

Nonis, S., & Hudson, G. (2006). Academic Performance of College Students: Influence of Time Spent Studying and Working. *The Journal of Education for Business, 81*, 151–159. https://doi.org/10.3200/JOEB.81.3.151-159

Pearson, K. (1902). Mathematical contributions to the theory of evolution. XI.-On the influence of natural selection on the variability and correlation of organs. *Proceedings of the Royal Society of London, 69*(451–458), 330–333.

Pek, J., Wong, O., & Wong, A. (2018). How to Address Non-normality: A Taxonomy of Approaches, Reviewed, and Illustrated. *Frontiers in Psychology, 9*, 2104.

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology, 108*(4), 455.

Petz, B., Kolesarić, V., & Ivanec, D. (2012). *Petz's statistics. Basis statistics methods for nonmathematicians*. Naklada Slap.

Pintrich, P. R. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. NCRIPTAL, The University of Michigan.

Plant, E. A., Ericsson, K. A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade point average across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology, 30*(1), 96–116.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322–338.

Primi, R., Ferrão, M. E., & Almeida, L. S. (2010). Fluid intelligence as a predictor of learning: A longitudinal multilevel approach applied to math. *Learning and Individual Differences, 20*(5), 446–451.

Rau, W., & Durand, A. (2000). The academic ethic and college grades: Does hard work help students to "make the grade"? *Sociology of Education, 73*, 19–38.

Raven, J. (1991). *Court, JH (2003, updated 2004). Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio TX: Harcourt Assessment.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28*, 164–171.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353–387.

Rimfeld, K., Kovas, Y., Dale, P. S., & Plomin, R. (2016). True grit and genetics: Predicting academic achievement from personality. *Journal of Personality and Social Psychology, 111*(5), 780–789.

Roberts, M. C., Meier, R. S., Santogrossi, D. A., & Moore, D. R. (1978). Relationship of student characteristics and performance in a personalized system of instruction course. *Teaching of Psychology, 5*(3), 118–121.

Rosário, P., Núñez, J. C., Valle, A., González-Pienda, J., & Lourenço, A. (2013). Grade level, study time, and grade retention and their effects on motivation, self-regulated learning strategies, and mathematics achievement: A structural equation model. *European Journal of Psychology of Education, 28*(4), 1311–1331.

Sadler, P. M., & Tai, R. H. (2001). Success in introductory college physics: The role of high school preparation. *Science Education, 85*(2), 111–136.

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin, 143*(6), 565–600.

Schuman, H., Walsh, E., Olson, C., & Etheridge, B. (1985). Effort and reward: The assumption that college grades are affected by quantity of study. *Social Forces, 63*(4), 945–966.

Schwarz, N., & Sudman, S. (2012). *Autobiographical memory and the validity of retrospective reports*. Springer Science & Business Media.

Shapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal, 41*(1), 159–189.

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review, 76*(2), 179–193. https://doi.org/10.1037/h0027277

Shipley, B. (2016). *Cause and correlation in biology: A user's guide to path analysis, structural equations and causal inference with R*. Cambridge University Press.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science, 11*(6), 771–799.

Shum, M. S. (1998). The role of temporal landmarks in autobiographical memory processes. *Psychological Bulletin, 124*(3), 423–442.

Smid, S. C., McNeish, D., Miočević, M., & van de Schoot, R. (2020). Bayesian versus frequentist estimation for structural equation

models in small sample contexts: A systematic review. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(1), 131–161.

Smith, P. (1999). *Whetton C. Testovi Općih Sposobnosti: Priručnik Za Primjenu i Bodovanje*. Naklada Slap.

Soares, D. L., Lemos, G. C., Primi, R., & Almeida, L. S. (2015). The relationship between intelligence and academic achievement throughout middle school: The role of students' prior academic performance. *Learning and Individual Differences, 41*, 73–78.

Spearman, C. (1904). ' General Intelligence', objectively determined and measured. *The American Journal of Psychology, 15*(2), 201–292.

Spearman, C. E. (1927). *The abilities of man (Vol. 89)*. Macmillan New York.

Stauffer, J. M., Ree, M. J., & Carretta, T. R. (1996). Cognitive-components tests are not much more than g: An extension of Kyllonen's analyses. *The Journal of General Psychology, 123*(3), 193–205.

Sudman, S., & Bradburn, N. M. (1982). *A practical guide to questionnaire design*. Jossey-Bass.

Swanson, H. L., & Alloway, T. P. (2012). *Working memory, learning, and academic achievement In APA educational psychology handbook, Vol 1 Theories, constructs, and critical issues. (pp 327–366)*. American Psychological Association.

Tai, R. H., Sadler, P. M., & Loehr, J. F. (2005). Factors influencing success in introductory college chemistry. *Journal of Research in Science Teaching: THe Official Journal of the National Association for Research in Science Teaching, 42*(9), 987–1012.

Tetzner, J., Becker, M., & Maaz, K. (2017). Development in multiple areas of life in adolescence: Interrelations between academic achievement, perceived peer acceptance, and self-esteem. *International Journal of Behavioral Development, 41*(6), 704–713. https://doi.org/10.1177/0165025416664432

Unsworth, N. (2016). Working memory capacity and recall from long-term memory: Examining the influences of encoding strategies, study time allocation, search efficiency, and monitoring abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(1), 50–61.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*(1), 104–132.

Vaci, N., Gula, B., & Bilalić, M. (2014). Restricting range restricts conclusions. *Frontiers in Psychology, 5*, 569.

Vaci, N., Cocić, D., Gula, B., & Bilalić, M. (2019a). Large data and Bayesian modeling—aging curves of NBA players. *Behavior Research Methods, 51*(4), 1544–1564.

Vaci, N., Edelsbrunner, P., Stern, E., Neubauer, A., Bilalić, M., & Grabner, R. H. (2019b). The joint influence of intelligence and practice on skill development throughout the life span. *Proceedings of the National Academy of Sciences, 116*(37), 18363–18369. https://doi.org/10.1073/pnas.1819086116.

Van Buuren, S. (2018). *Flexible imputation of missing data*. CRC Press.

Vedel, A. (2014). The Big Five and tertiary academic performance: A systematic review and meta-analysis. *Personality and Individual Differences, 71*, 66–76.

von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The Hungry Mind: Intellectual Curiosity Is the Third Pillar of Academic Performance. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science, 6*(6), 574–588. https://doi.org/10.1177/1745691611421204

Wachs, M. C., & Harris, M. (1986). Simultaneous and successive processing in university students: Contribution to academic performance. *Journal of Psychoeducational Assessment, 4*(2), 103–112.

Wang, D., Yuan, F., & Wang, Y. (2022). Growth mindset and academic achievement in Chinese adolescents: A moderated mediation model of reasoning ability and self-affirmation. *Current Psychology, 41*(2), 783–792. https://doi.org/10.1007/s12144-019-00597-z

Yeager, D. S., & Dweck, C. S. (2020). What can be learned from growth mindset controversies? *American Psychologist, 75*(9), 1269.

Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., Lee, H. Y., O'Brien, J., Flint, K., & Roberts, A. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of Educational Psychology, 108*(3), 374–391.

Yuan, K., Steedle, J., Shavelson, R., Alonzo, A., & Oppezzo, M. (2006). Working memory, fluid intelligence, and science learning. *Educational Research Review, 1*(2), 83–98.

Zaboski, B. A., II., Kranzler, J. H., & Gage, N. A. (2018). Meta-analysis of the relationship between academic achievement and broad abilities of the Cattell-Horn-Carroll theory. *Journal of School Psychology, 71*, 42–56.

Zhang, Z. (2014). Monte Carlo based statistical power analysis for mediation models: Methods and software. *Behavior Research Methods, 46*(4), 1184–1198.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166–183.

Zondervan-Zwijnenburg, M., Peeters, M., Depaoli, S., & Van de Schoot, R. (2017). Where do priors come from? Applying guidelines to construct informative priors in small sample research. *Research in Human Development, 14*(4), 305–320.