



Formal and Informal Learning and First-Year Psychology Students' Development of Scientific Thinking: A Two-Wave Panel Study

Demet Soyylmaz¹, Laura M. Griffin², Miguel H. Martín^{3,4}, Šimon Kucharský⁵, Ekaterina D. Peycheva⁶, Nina Vaupotič⁷ and Peter A. Edelsbrunner^{8*}

¹ Department of Psychology, Istanbul Bilgi University, Istanbul, Turkey, ² Faculty of Film, Art and Creative Technologies, Dún Laoghaire Institute of Art, Design and Technology, Dún Laoghaire, Ireland, ³ Faculty of Psychology, Pontifical University of Salamanca, Salamanca, Spain, ⁴ Faculty of Psychology, Ghent University, Ghent, Belgium, ⁵ Department of Psychology, University of Amsterdam, Amsterdam, Netherlands, ⁶ Department of General, Experimental and Genetic Psychology, Sofia University St. Kliment Ohridski, Sofia, Bulgaria, ⁷ Department of Psychology, University of Ljubljana, Ljubljana, Slovenia, ⁸ Research on Learning and Instruction, Department of Humanities, Social and Political Sciences, ETH Zurich, Zurich, Switzerland

OPEN ACCESS

Edited by:

Kristina Egumenovska,
International School for Advanced
Studies, Italy

Reviewed by:

Tom Rosman,
Leibniz Institute for Psychology
Information, Germany
Caitlin Drummond,
Carnegie Mellon University, USA

*Correspondence:

Peter A. Edelsbrunner
peter.edelsbrunner@ifv.gess.ethz.ch

Specialty section:

This article was submitted to
Quantitative Psychology
and Measurement,
a section of the journal
Frontiers in Psychology

Received: 07 November 2016

Accepted: 18 January 2017

Published: 10 February 2017

Citation:

Soyylmaz D, Griffin LM, Martín MH,
Kucharský Š, Peycheva ED,
Vaupotič N and Edelsbrunner PA
(2017) Formal and Informal Learning
and First-Year Psychology Students'
Development of Scientific Thinking:
A Two-Wave Panel Study.
Front. Psychol. 8:133.
doi: 10.3389/fpsyg.2017.00133

Scientific thinking is a predicate for scientific inquiry, and thus important to develop early in psychology students as potential future researchers. The present research is aimed at fathoming the contributions of formal and informal learning experiences to psychology students' development of scientific thinking during their 1st-year of study. We hypothesize that informal experiences are relevant beyond formal experiences. First-year psychology student cohorts from various European countries will be assessed at the beginning and again at the end of the second semester. Assessments of scientific thinking will include scientific reasoning skills, the understanding of basic statistics concepts, and epistemic cognition. Formal learning experiences will include engagement in academic activities which are guided by university authorities. Informal learning experiences will include non-compulsory, self-guided learning experiences. Formal and informal experiences will be assessed with a newly developed survey. As dispositional predictors, students' need for cognition and self-efficacy in psychological science will be assessed. In a structural equation model, students' learning experiences and personal dispositions will be examined as predictors of their development of scientific thinking. Commonalities and differences in predictive weights across universities will be tested. The project is aimed at contributing information for designing university environments to optimize the development of students' scientific thinking.

Keywords: epistemic cognition, informal learning, need for cognition, self-efficacy, scientific thinking, psychology students

INTRODUCTION

Scientific thinking encompasses purposeful thinking with the aim to enhance knowledge, using the abilities to generate, test and revise theories as well as being able to reflect on how knowledge is acquired and changed (Kuhn, 2002). It is a prerequisite for engagement in scientific inquiry (Kuhn et al., 2000). Little is known about the development of scientific thinking during aspiring researchers' development, especially at the early undergraduate level (Parent and Oliver, 2015).

The 1st-year at university is a particularly critical educational period for students' development of skills, interests, and aspirations (Jenert et al., 2015). For those students who want to become researchers, 1st-year education gives a first impression of science, with core courses found in psychology programs such as research methods and statistics (Stoloff et al., 2009). These educational experiences contribute to students' future motivation and understanding of science (Manning et al., 2006; Jenert et al., 2015). Not all students know in their 1st-year of university education if they want to engage in research and become researchers. But if they decide so at any point in the future, it is necessary to equip them with advanced scientific thinking as early as possible. Thus, the quality of 1st-year education can influence students' further aspirations and development as students and researchers. In this study, we examine the influence of learning experiences in 1st-year psychology on students' development of scientific thinking. One aim of the present study is to identify learning experiences related to the development of scientific thinking in the 1st-year of higher education as well as to pinpoint those that are most prevalent among successful scientific thinkers. This way, we try to capture the experiential profile of budding psychology researchers. Such findings could be of vital service in the development of psychology curricula that better reflect the learning needs of aspiring researchers as well as motivate students to become such.

A facet of scientific thinking that we consider central for potential future researchers is scientific reasoning¹. Scientific reasoning delineates the skills needed to conduct scientific inquiry, such as argumentation, drawing inferences from data, and engaging in experimentation (Zimmerman, 2007). It includes understanding and identifying relevant variables and how to interpret the information obtained from an experiment and various other research designs (Klahr, 2000).

Related to scientific reasoning skills, students should understand basic statistical concepts to evaluate the strength and uncertainty of scientific evidence. Previous research has shown that misconceptions about common statistical indicators, such as *p*-values and confidence intervals, are prevalent in student and teacher populations (Hoekstra et al., 2014; Morey et al., 2015), and in published research literature (Gelman and Stern, 2006; Nieuwenhuis et al., 2011). It is not surprising that students develop misconceptions about statistics, given that classical statistical methods violate common sense (Wagenmakers, 2007; Duffy, 2010). Psychology students commonly learn about null hypothesis significance testing in their statistics courses, which leads them to make interpretations based on arbitrary *p*-values previously set by researchers. This method violates common sense and leads to misunderstandings because the results are actually based on observations that have not occurred (Wagenmakers, 2007). For example, researchers intuitively tend to think that hypotheses tests inform them about the probability that the alternative hypothesis is true, but hypotheses

tests based on *p*-values have a different aim; they only inform about the long-term frequency of possible data given the null hypothesis. Crucially, those misconceptions can lead to wrong inferences, both in the conduct of research but also in the evaluation of published literature, which might have contributed to a current crisis in the confidence of psychological science (Pashler and Wagenmakers, 2012). With this in mind, we regard the absence of statistics misconceptions as a relevant aspect of scientific thinking in today's psychology students as developing researchers.

Another core facet of scientific thinking is epistemic cognition, which encompasses beliefs about knowledge, knowing, and the processes by which those beliefs are formed and influence further learning (Hofer and Pintrich, 1997; Kitchener, 2002; Greene et al., 2008). Greene et al. (2010) developed a model of epistemic and ontological cognition that integrates prior models by positing positions and dimensions. Each of four positions (Realism, Dogmatism, Skepticism, and Rationalism) corresponds to a distinctive pattern of individuals' beliefs along the three dimensions of simplicity and certainty, justification by authority, and personal justification of knowledge (Hofer, 2000). Simple and certain knowledge refers to the opinion that knowledge is isolated, simple and constant over time, justification by authority reflects a belief that knowledge can be ambiguous but holds greater weight when presented by an authority figure, and personal justification is a belief that all information presented must be engaged with critically before judging it to be true, and even then, it may not remain true over time. The first developmental stage, realism, represents strong beliefs in simple and certain knowledge, justification by authority, and personal justification. The position of dogmatism is demonstrated through strong justification by authority. The position of skepticism reflects strong personal justification. Lastly, rationalism presents moderate agreement with justification by authority and personal justification, but strong disagreement with the concept of simple and certain knowledge. Epistemic cognition can influence critical thinking, scientific argumentation, and learning (Kuhn et al., 2000; Nussbaum et al., 2008; Franco et al., 2012), and it is related to students' self-regulated learning (Bråten and Strømsø, 2005). Several studies thus far have linked epistemic cognition with students' learning achievements. A study conducted by Muis and Franco (2009) illustrates that epistemic cognition directly influences achievement goals of students in educational psychology course, which, in turn influenced their engagement in their tasks and final course achievement. In the similar vein, Bråten and Ferguson (2014) showed that epistemic beliefs contribute to achievement over and above cognitive capacity and personality traits of students. Moreover, Chen et al. (2014) showed that students who are self-efficacious about learning science, approach a task by examining arguments from several sources to make a decision, thus indicating a moderating role of self-efficacy in how epistemic cognition is related to academic outcomes. The results of these studies therefore support the idea that epistemic cognition plays a major role in students' further engagement and development of scientific thinking and we consider it a facet of scientific thinking that should develop early in psychology students.

¹Scientific reasoning is sometimes used interchangeably with scientific thinking (and sometimes with the related but broader construct critical thinking, Halpern et al., 2012) but in the context of the present research we deem it useful to conceptualize scientific reasoning and epistemic beliefs as facets of scientific thinking.

The development of scientific thinking begins in the first stages of life and continues throughout childhood and adolescence up into adulthood (Sodian et al., 1991; Kuhn et al., 2000, 2015; Zimmerman, 2007). However, this development does not occur automatically but through steady input from deliberate learning experiences (Kuhn, 2002, 2009; Klahr et al., 2011). Perhaps the most conspicuous way of improving the level of scientific reasoning is through formal education. In fact, research shows that students with higher level of education as well as students who were exposed to a research methodology course are more likely to exhibit better scientific reasoning skills (Lehman and Nisbett, 1990; Lawson et al., 2000). Demonstrating and engaging students in quantitative methodology and scientific inference improves their skills in theoretical modeling and experimentation (Duschl and Grandy, 2013; Holmes et al., 2015). Further learning experiences that predict university students' scientific thinking include collaborative learning (e.g., Gokhale, 1995; Johnson et al., 1998), social media use (Ebner et al., 2010; Dabbagh and Kitsantas, 2012; Kassens-Noor, 2012; Vivian et al., 2014), taking research methods and statistics courses (Lehman and Nisbett, 1990; VanderStoep and Shaughnessy, 1997), passive and active participation in research projects (Wayment and Dickson, 2008; VanWormer et al., 2014) and taking laboratory modules that include interpreting the results of an experiment (Coleman et al., 2015). Thus, university environments offer varied experiences that can help students develop scientific thinking.

In order for students to develop scientific thinking, it is not sufficient that relevant learning opportunities are offered at university. It is necessary that students show high engagement in formal activities and beyond. In the current study, we therefore look into students' engagement in relevant formal and informal learning activities. Informal learning at university can be defined as self-directed learning in the sense that the student chooses the topic, curriculum, and contents, and learning and assessment modalities, with the aim to develop knowledge, skills, or competences (Hofstein and Rosenfeld, 1996; Laurillard, 2009). It is closely related to conceptions of student engagement (Krause and Coates, 2008), as well as self-sustained, self-initiated, and free choice learning (Falk, 2001; Barron, 2006; Yang, 2015). Based on this definition, examples of informal learning experiences are attending science conferences, reading scientific books, and engaging in science-related discussions with peers. Formal learning, in comparison, is highly structured through university bodies in its curriculum, fixed learning activities, and assessment, with a course achievement or qualification as an end product (Resnick, 1987; Eshach, 2007; Patrick, 2010). This distinction posits informal learning as interest-driven, in comparison to formal learning as curriculum-based, assessment-driven, and qualification-oriented activities.

What factors predict students' engagement in learning activities that are likely to foster their scientific thinking? A characteristic that we take into account is students' science self-efficacy, that is, the confidence they have in their ability to do science (Beißert et al., 2014). Self-efficacy, defined as the belief in own capability to succeed (Bandura, 1997), is a major predictors of university students' cognitive engagement,

academic persistence in science-related courses, and career choices (Chemers et al., 2001; Walker et al., 2006; Chen and Usher, 2013). Along with self-efficacy, need for cognition has been shown to predict academic success (Elias and Loomis, 2002). It is a stable tendency to engage in and enjoy effortful thinking (Cacioppo and Petty, 1982). Need for cognition is related to intellectual engagement and positive attitudes toward effortful tasks, and thus with a richer personal history of gaining knowledge on a variety of topics (Woo et al., 2007).

In the current study, we assess psychology students twice to examine the contributions of formal and informal learning experiences to their development of scientific reasoning including statistics misconceptions, and epistemic cognition during the 1st-year of study. The assessments take place at the beginning and again at the end of their second semester. We aim to examine interrelations in the development of scientific reasoning and epistemic cognition during the semester, and the contribution of students' engagement in both types of learning experiences to this development. This includes their additive effects, and the involvement of students' self-efficacy and need for cognition in these effects. Our definition of informal learning posits that it is self-guided and goes beyond the mere aim of finishing courses and obtaining grades. It thus related informal learning strongly with interest-driven student engagement. Student engagement in educationally purposeful activities is positively related to academic outcomes in 1st-year students as well as students' persistence at the same institution (Kuh et al., 2008). Similarly, student engagement has been linked to desirable learning outcomes such as critical thinking and academic achievement (Carini et al., 2006). Whereas engagement in formal learning could stem from internal or external motivational factors, engagement in informal activities represents only intrinsically motivated behavior, which is derived from interest and performed for pleasure and desire. For these reasons we assume that informal learning contributes to students' development of scientific reasoning and epistemic cognition, beyond formal learning. The overarching aim of the design is to establish the circumstances under which potential future researchers in psychological science are able to develop scientific thinking during the early stages of their studies. We therefore examine specific patterns of scientific thinking and its predictors in students who identify themselves as aspiring researchers.

MATERIALS AND METHODS

Design and Sample

The study has a two-wave correlational panel design. Participants will be drawn from the 1st-year psychology courses of 11 universities from eight countries across Europe. We collaborate with 1st-year professors from each university. The choice of universities was based on personal affiliations and on the aim of gathering students from diverse backgrounds across Europe. Psychology student cohorts at the universities span between 40 and 700 students. Students from eight universities will participate during a regular class lesson and the remaining three

universities online. At three of the universities, students will receive assessment credits for participating in the study.

Sample size planning based on power analysis is not relevant because we will use Bayesian estimation and hypothesis testing for statistical analysis (Etz et al., 2016; Wagenmakers et al., 2016). In this statistical framework, power is not conceptualized because hypothesis testing is not based on an inferential framework but on continuous evaluation of evidence (Schönbrodt and Wagenmakers, 2016).

Materials and Equipment

Choice of Measures

For every construct that we aim to assess, a literature search was done in the PsycInfo and Scopus databases to identify available measures. The choice of the instruments was based on psychometric quality, appropriateness for university context, administration time, translation feasibility, and meaningfulness of usage in a variety of international universities. Regarding psychometric quality, we ensured that basic analysis such as factor analysis, estimation of reliability or internal consistency had been conducted and achieved at least moderate results. Appropriateness for the 1st-year of university was taken into account insofar as we tried to estimate on which level Psychology students develop during their 1st-year. For example, scientific reasoning is a broad construct, and we chose an instrument that assesses skills which we think are critical for students' further development, and likely to show at least some development already during their first university year. The chosen instrument assesses principles of experimental design that we deem relevant for understanding the critical quality characteristics of any research the students learn about (Drummond and Fischhoff, 2015).

Demographics Questionnaire

Students' demographic characteristics will include their age, gender, former university education, career aspirations, grades in high school, the grade of first university examination and family socioeconomic status (see Appendix A). For the latter, we will ask students about their parents' highest achieved education, bedroom availability and the number of books at home in their adolescence (Evans et al., 2010). Socioeconomic status is assessed to examine its influence on the main study variables and to estimate other variables' influence while controlling for it. We will assess family socioeconomic status because university students are still in education, which constrains their own educational level and also their working situation, the most common indicators of personal socioeconomic status. Family socioeconomic status is thus commonly assessed for research in academic contexts (Caro and Cortes, 2012). Students' estimated score from the first principal component of the four variables will be used as an indicator of their family socioeconomic status. Finally, we will assess the quantity of formal education relevant for developing scientific thinking (number of methodology and statistics-related courses, number of philosophy of science and epistemology-related courses).

Scientific Reasoning

As a measure of scientific reasoning, the *Scientific Reasoning Scale* developed and validated by Drummond and Fischhoff (2015) will be used. It contains eleven true or false items in which hypothetical research scenarios are described and the participant has to decide whether the scenario can lead to proposed inferences. Each of the items relates to a specific concept crucial for the ability to come to valid scientific conclusions. The concepts include understanding the importance of control groups and random assignment, identifying confounding variables, and distinguishing between correlation and causation. Scores on the SRS show adequate internal consistency (Cronbach's $\alpha = 0.70$) and correlate positively with cognitive reflection, numeracy, open minded thinking, and the ability to analyze scientific information (Drummond and Fischhoff, 2015). Following this scale, we added an additional item assessing students' understanding of sample representativeness (Appendix B). Students' mean score on the scale will be used in descriptive analysis as an indicator of their scientific reasoning. Whether the item on sample representativeness can be added to the scale will be decided based on a confirmatory factor analysis: It will be added in case its factor loading is within the range of the other items.

Statistics Misconceptions

We developed a questionnaire encompassing five questions that deal with common statistical misconceptions (Appendix B). Items dealing with p -value and confidence interval misinterpretations were taken directly from Gigerenzer (2004) and Morey et al. (2015). We chose the item with the highest prevalence of wrong answers among university students from each article to achieve high variance in our sample of 1st-year students. We further developed items similar in structure dealing with the interpretation of non-significant results, the equivalence of significant and non-significant results (Gelman and Stern, 2006; Nieuwenhuis et al., 2011), and sample representativeness. The items share structure and answer format with the scientific reasoning scale by Drummond and Fischhoff (2015). We added the items after the end of the scientific reasoning scale. Participants are also asked whether they have ever learned about p -values, confidence intervals, and sample representativeness. In case they check "no," their answers on the respective questions will be treated as missing values. Students' mean value across the four questions dealing with p -values and confidence intervals will be used as an indicator of their statistics misconceptions. The question on sample representativeness, as described above, will be used as an additional item of the scientific reasoning scale.

Validation Questions

For the Scientific Reasoning Scale and the added statistics misconceptions items, we will add one open-answer validation question. Each student will receive the following question at one random item of the 16 items that the two scales encompass: "Why did you choose this answer? Please provide an explanation.", followed by two lines on which the students are supposed to provide a short rationale for their multiple choice-answer. The question to which this additional open answer is added will

differ randomly between students, so that a random subsample of the students will deal as validation sample for each question. We implement this validation measure because the SRS to the best of our knowledge has not yet been translated into our sampled languages and not been used in the sampled countries. It is therefore necessary to examine whether 1st-year psychology students' answers on these questions reflect the target construct. The statistics misconceptions to the best of our knowledge have not yet been thoroughly validated but rather used to assess the prevalence of wrong answers among students and academics, and we developed three of the questions on our own, therefore we include them in this validation procedure.

Epistemic Cognition

To assess epistemic cognition we will administer the Epistemic and Ontological Cognition Questionnaire (EOCQ; Greene et al., 2010). It contains 13 items and a 6-point item response scale ranging from 1 (completely disagree) to 6 (completely agree). The instrument takes into account the contextuality of epistemic cognition by providing the opportunity to insert a domain into the item stems (Greene et al., 2008). We insert *Psychology* and *Psychological science* for the domain that the students should rate the items about. Five items represent simple and certain knowledge (example: “*in psychological science, what is a fact today will be a fact tomorrow*”), four items represent justification by authority (“*I believe everything I learn in psychology class*”), and four items represent personal justification (“*in psychological science, what's a fact depends upon a person's view*”). Higher ratings of ten items indicate stronger beliefs and high ratings of three items indicate weaker beliefs. Reliability estimates (*H* coefficient) range from 0.45 to 0.90 depending on facet and context (Greene et al., 2010). Mean scores on all three subscales will undergo mixture modeling analysis, which will yield an epistemic cognition-profile for each student that will be used for further analysis (Greene et al., 2010).

Need for Cognition

We will use the Need for Cognition Short Scale (NFC-K; BeiBert et al., 2014) to measure the tendency to engage in and enjoy thinking. The short scale is a modified 4-item version of the 18-item Need for Cognition Scale created by Cacioppo and Petty (1982). On a 7-point scale the students are asked to rate to which extent they agree with four simple statements. An example item is “*I would prefer complex to simple problems.*” Mean scores from this scale will be used for descriptive analysis, with higher scores indicating that students are more motivated to apply their thinking skills. Test retest reliability is $r = 0.78$, Cronbach's $\alpha = 0.86$ (BeiBert et al., 2014). The score will be used to predict students' development of scientific thinking, and also as a control variable to examine which variables predict students' development beyond need for cognition.

Science Self-Efficacy

The Science Self-Efficacy (SSE) scale, which consists of 10-items used by Moss (2012) will be used (Cronbach's $\alpha > 0.80$). It is a modified version of a vocational self-efficacy survey designed by Riggs et al. (1994). It particularly aims to measure confidence in

skills to engage in scientific inquiry. The items are rated on a scale from 1 to 10 (1 = *not able or not true at all*, 10 = *completely able or completely true*). An example item is “*I have all the skills needed to perform science tasks very well.*” Students' mean score on the scale will be used for statistical modeling. The score will be used to predict students' development of scientific thinking, and also as a control variable to examine which variables predict students' development beyond science self-efficacy.

Formal and Informal Learning Experiences

We developed a survey to assess students' engagement in learning experiences that we presume relevant for the development of scientific thinking (Appendix C). The selection of experiences is based on the discussed literature, and it will be further informed and adapted based on the pilot study interviews (Appendix D). Our definitions of formal and informal learning imply a continuum of formality within and across learning activities. For example, a frequent formal learning activity is the studying of a text that is mandatory reading for a research methods course. When students gain interest in the text contents, they might initiate further voluntary reading to inform themselves beyond the course requirements, which in our definition is then an informal learning experience. Our assessment method encompasses a wide variety of prescribed and non-prescribed scientific learning experiences: For each of the assessed activities that can be either formal or informal, we ask students how often they engaged in these as part of mandatory course activities, or for reasons going beyond these. Specifically, we let students rate subjectively for experiences where this applied how much they engaged in them because it was obligatory for course requirements (formal engagement), because it was obligatory but they were also interested (formal and informal engagement), or merely out of own interest (informal engagement).

In the second part of the survey, we ask students about the most relevant three courses they took that were related to research methods, statistics, science, history of science or other similar concepts. We ask for up to three courses because we studied the official bachelor curricula from the targeted universities and most students will not have more highly relevant courses during their first and second semester. Therefore, reporting on further courses might make it strongly subjective which courses the students deem relevant to this question, and it might take rather long and be exhausting to report details on any relevant courses they could think of. To check that they did not have many more relevant courses we, however, ask in the demographics for the absolute numbers of relevant courses. Thus, for up to three most relevant courses, they first list the names of the courses and whether the courses were mandatory or elective. Then, we ask students about their general engagement in these courses (student presence, devoted working time), and course quality (ratings of overall course quality, teaching quality, frequency of inquiry and reflective course elements). Finally, reflecting informal engagement, they rate how much they engaged in each of these courses out of their motivation or interest, beyond the course requirements. Estimating principal components, we will weigh general course engagement across courses with course content ratings to yield an indicator of formal engagement,

and informal (out of own motivation or interest) engagement with course quality ratings to yield an indicator of informal engagement.

Translations and Pilot Study

Considering students from different countries' levels of competence in English may not be sufficiently high, the materials and instruments have been translated into Spanish, Slovenian, Turkish, Bulgarian, and Czech by the local researchers from these countries. Then, they have been back translated by bilingual speakers to enable reconciliation of the translated texts with the original. During a pilot study, the materials and instruments were administered to a small number (10 from each country) of 2nd–4th year psychology students with cognitive surveying and interviewing to identify problematic passages in terms of ambiguous or confusing instructions and translations (Ziegler et al., 2015). During the cognitive surveying, participants were asked to read the instructions and items aloud. After each passage, they were instructed to report everything that came to their mind when thinking about the instruction or item and what they were thinking while answering the items. In the end, they were asked to reflect freely on the purpose, comprehensibility, and quality of the instrument. This data were used to alter potential problematic passages. Proposed changes were again translated back to English for comparison with the original. Pilot participants were also requested to respond to several interview questions regarding their formal and informal educational experiences throughout their lives that they believe might have contributed to their scientific reasoning and epistemic cognition (Appendix D). Their responses served to improve the formal and informal experiences survey, so that it would more adequately reflect students' relevant learning experiences.

Stepwise Procedures

The data will be collected at two time points. The first assessment will be conducted during the first 2 weeks of the second semester (between January and March) and the second will take place within the last 2 weeks of the academic schedule before exams (May and June), depending on each university's calendar. For universities at which collaborators agree to in-class assessments, these will take place directly in the classrooms or other provided university space. Professors will be asked to reward students with course credits for research participation, depending on the ethical policy of the institution. Ideally, with the professors' prior consent, the entire first-year courses will be assessed during a lecture. The local researchers in each country will distribute the questionnaires before the assessment starts and collect them afterward. In case an in-class administration of our instruments is not possible, we will ask the students to participate in an online version of our assessments. An online version has been prepared in the Qualtrics (Qualtrics, Provo, UT, USA) environment with a similar structure to the pen and paper version. For the online version, students will be provided with a hyperlink and encouraged to fill it in at their convenience within a week. If they have not yet finished the survey, they will receive a reminder email 2 days before this time limit.

The questionnaires will have the same structure in the in-class and in the online version. In both cases, participants will be given a short explanation of what the research is for, and what their participation will entail, which will be read aloud by the experimenter in class. They will then be asked to read an information sheet and read and sign a consent form, prior to proceeding with the assessment. One administration process is expected to last for about 35 min. The scales more strongly related to cognitive skills will be presented in the beginning of the assessment and the learning experiences will be assessed in the end to prevent the experiential questions from influencing later answers. The structure of the assessments is depicted in **Figure 1**.

The participants will first be asked to compose an identification code consisting of their mother's and father's initials, and the month that they were born (in mm format). They will then be asked to complete the demographics information about themselves. They will subsequently proceed to complete five scales measuring scientific reasoning, statistical misconceptions, epistemic cognition, need for cognition, and scientific self-efficacy. This will reflect their skills and attitudes after one semester studying psychology. In addition, they will be asked to complete the survey regarding their formal and informal learning experiences during the first study semester. In the end of the assessment, the students will be thanked for their participation, and informed that the assessment will take place again in the end of the semester. They will be asked to refrain from discussing the assessment with each other or to look up the contents.

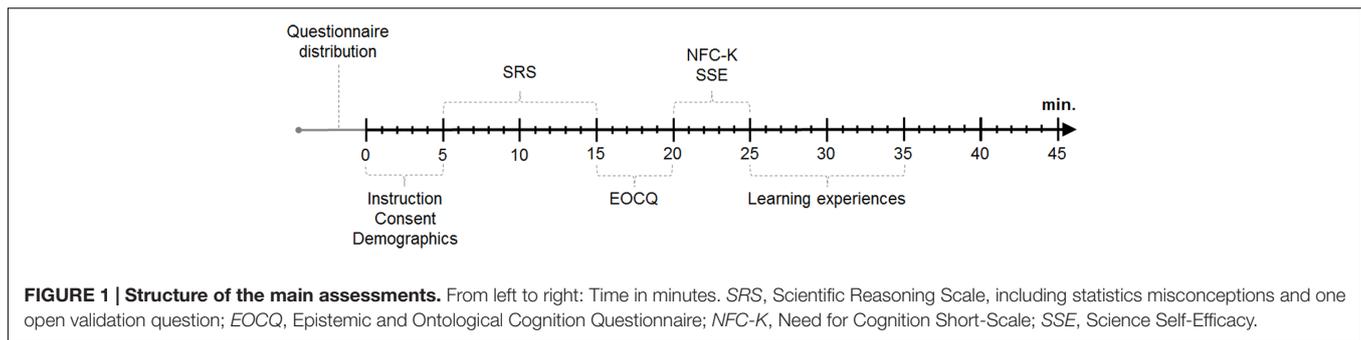
At the end of the second semester, participants will be approached as before to participate in the study. They will be asked to write their identification number as before, and to again complete the same scales as at the first assessment. On the demographics sheet, this time they will be asked additionally whether and to which extent they discussed the contents of the first assessment with peers or looked up the contents. The survey on formal and informal learning experiences will this time be referring to experiences during their second study semester. At this point participants will be given a debriefing form and thanked for their participation.

PROPOSED ANALYSIS AND ANTICIPATED RESULTS

Qualitative Data Analysis

Qualitative data will stem from the pilot study interviews. Transcriptions of the interviews will be analyzed using content analysis. The analysis will be aimed at identifying relevant formal and informal learning experiences else than those known from available literature. Insights from these data will be used to refine the learning experiences questionnaire for the main assessments.

We will also analyze the open validation questions about scientific reasoning and statistical misconceptions to see whether the correct multiple choice-answers on the items reflect the intended concepts (Drummond and Fischhoff, 2015). Given the fact that the items are forced choice between true or false, we of course expect that some of the correct answers will be a



result of guessing. We will encode whether the rationale for the answer is sufficient to come to the correct choice given the specific question. Then, we will try to group answers that did not have the correct rationale for the answer to see common misunderstandings. This will inform us about the validity of the items. We will also seek for common misconceptions leading to erroneous answers and try to categorize them with open followed by axial coding to get insight into why students make mistakes regarding the specific concepts. This will inform us what aspects of given concepts are hard to grasp and which misconceptions should be deliberately targeted by university lecturers.

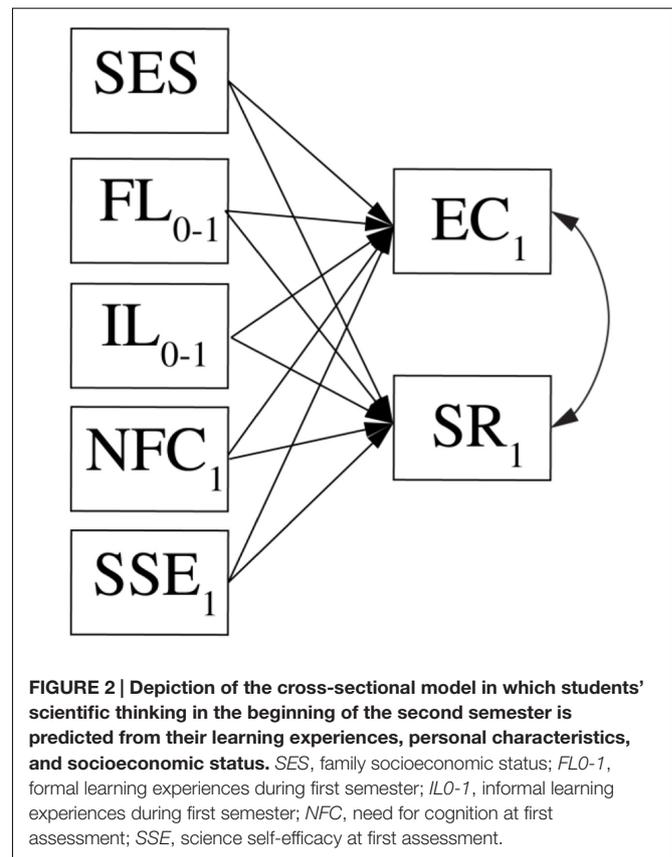
Confirmatory Statistical Analysis

Bayesian structural equation modeling will be applied to examine our main research questions. The models will be written using the *r2jags* (Su and Yajima, 2012) and *rjags* (Plummer, 2013) packages in the R software (R Core Team, 2013) to be estimated in the JAGS software (Plummer, 2003). There will be two main models. In a cross-sectional model, we will predict students' scientific thinking and research aspirations after the first half year from their educational experiences during the first half year, science self-efficacy and need for cognition, and from their learning experiences and family socioeconomic status. A depiction of the structural relations in this model is provided in **Figure 2**.

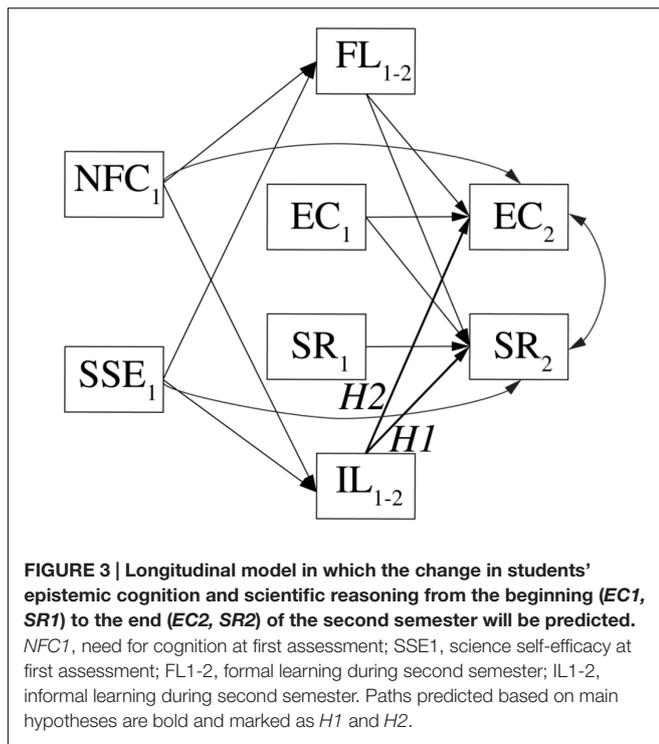
In a longitudinal model, we will examine developmental interrelations between students' scientific reasoning and epistemic cognition, and predict their development from students' formal and informal learning experiences during the second half year, and how these are influenced by students' personal characteristics. A depiction of the structural relations in this model is provided in **Figure 3**.

Bayes factors will be used for hypothesis testing. The tested predictions based on our hypotheses include that informal learning experiences predict the development of scientific reasoning and epistemic cognition, controlling for formal learning experiences, need for cognition, and science self-efficacy.

The models will be estimated separately for students from each university to examine commonalities and differences in predictive weights. Since sample sizes vary strongly between countries and institutions, samples from some universities might not be sufficiently big to ensure convergence and precision of model estimation. Data from the biggest samples will therefore be used to slightly inform the parameter priors from the smaller samples. Only priors of non-focal (i.e., non-hypothesis-testing)



parameters will be informed in this way (for an overview of related techniques, see McElreath, 2016, pp. 424–430). This strategy is similar to hierarchical modeling but implies weaker partial pooling. Scripts for all confirmatory analyses will be uploaded to the Open Science Framework prior to data analysis. We will interpret the magnitude of obtained Bayes factors based on the accruing samples from the different universities. The Bayes factors will be computed as a ratio of likelihoods of two models that describe the theoretical alternatives we put to the test (Jarosz and Wiley, 2014). We derive those models for two focal hypotheses (depicted in the **Figure 3** as H1 and H2) here. Given we control for formal learning experiences (FL in **Figure 3**), need for cognition (NFC), and science self-efficacy (SSE), we expect that parameters for informal learning experiences (IL in **Figure 3**)



have a positive value. This expectation is equivalent to a one-sided hypothesis. We will therefore use so called one-sided hybrid Bayes factors (Morey and Rouder, 2011). As a null model, we will combine a point nil with part of the Cauchy distribution from the range of values of 0 up to the point where the effect size becomes important (equivalence region). As an alternative model, we will use the remaining of the Cauchy distribution. The Cauchy will have a scaling factor of 0.5, the equivalence region will be defined as $<0,0.1>$ and the mixture probability of the two parts of the null model will be equal to 0.5 (the point nil and equivalence region will have the same weights). This type of Bayes factors have been shown to possess desirable properties (Morey and Rouder, 2011). They asymptotically converge toward support for null or alternative if the true parameter lies in the area of one of the respective models, and remain indifferent if the true parameter lies on the boundary of null and alternative (0.1 in this case).

Missing data stemming from attrition or single non-answered items will be dealt with in the Bayesian analysis. Specifically, students' self-reported interest in research and becoming researchers and all other study variables that might be associated with participation willingness and missingness will be used to estimate students' missing data (see McElreath, 2016).

Exploratory Statistical Analysis

These analyses serve mainly to find specific patterns between all variables and trying to identify students interested in becoming future researchers. In addition, they will serve to develop potential hypotheses about profiles of psychology students interested in becoming researchers. For this research question, we do not have specific hypotheses and we will use exploratory analyses to examine how the intention to become a researcher

or not is associated with the other study variables. Specifically, we will use two methods. We will use network modeling to explore relations between the main study variables at the two assessments. For this analysis the *mgm*-package will be used, which can handle different distributions of the exponential family and applies regularization for sparse solutions (Haslbeck and Waldorp, 2016). The estimated networks at the two assessments will provide a concise and informative overview of interrelations between the study variables in the beginning and end of students' second semester. We will also estimate finite mixture models (Hickendorff et al., unpublished), to extract profiles of scientific thinking. We will examine profiles including scientific reasoning, the three epistemic cognition facets, and students' research aspirations as profile indicator variables, the two dispositional scales as profile predictors, and statistics misconceptions as a distal variable.

To further substantiate comparisons between universities and countries, we will examine measurement invariance, which shows whether the assessment instruments have comparable structure across different samples. Measurement invariance is not of critical importance for our hypotheses but it is informative for exploratory purposes, to see for example whether the instruments function similarly in the different languages. In our Bayesian framework, we will be able to handle small deviations from invariance by modeling approximate invariance (Van De Schoot et al., 2012).

LIMITATIONS

The design will allow estimating the predictive value of formal and informal learning experiences but causal conclusions are not fully warranted for various reasons. Controlling for students' general maturation in higher education would be possible by adding a control group, for example 1st-year students from a different field. Such a comparison would, however, be biased by self-selection effects because we cannot assign students randomly to different fields of study.

Another limitation is that measurement will take place twice within one semester, specifically within the second semester. This might be early for expecting students to develop in scientific reasoning, epistemic cognition, and also statistical misconceptions, depending on when these topics are part of students' courses. Not all students might learn about these topics in their 1st-year. We looked into students' official curricula at the target universities and in all of them there are courses that might be relevant but this is not clear for all universities. Regarding assessing twice within one semester, more change might be expected during a longer time period. We will, however, ask the students also for relevant experiences during the first semester, which overall will yield a picture of the whole 1st-year, which covers the focus of our study.

Also, because the collecting data will take place on two different occasions during the semester, there may be attrition between time one and time two. However, by aiming to collect data in class, we hope to maximize initial potential participation and minimize potential attrition rate by time two. Missing data

will also be minimized by ensuring that participants use the same identification code on both occasions. To maximize participation and minimize attrition, assessments will be conducted in-class at most universities. To minimize the missing data due to unintended skipping of the responses within the questionnaires, we will encourage students to thoroughly review their responses to ensure that they answered every question. In the online version, we will use automatized options to check missing responses to alert participants that they did not answer the question.

Another issue regards that we use self-report measures, particularly retrospective measures for students' learning experiences. These might be biased because remembering and subjectively judging the quality of courses from the last semester is error prone. Averaging ratings across students will hopefully lead to averaging out some error and the magnitude of between-student variance might indicate how error-prone retrospection is in this case. Also, for epistemic cognition it has been pointed out that self-report measures only allow quite superficial assessment (Mason, 2016). Alternatively, the idea of incorporating both quantitative and qualitative research methods upon EC has been supported (e.g., Greene and Yu, 2014). For the aims of our study we deem the EOCQ self-report measure appropriate but it will not allow a comprehensive look into the processes underlying students' development.

With reference to the country comparisons, even if it is not our main focus, they might be biased by the fact that at some universities all students participate within courses, at some they participate voluntarily without incentive, and at some voluntarily with incentive. Consequently, we acknowledge that these differences in recruitment may influence the results. In fact, voluntary participation in this research topic about scientific thinking can be seen as an indicator of student engagement and interest in science (and then, be categorized as an informal experience). In addition, it could happen that samples where students participate voluntarily with no incentive are overrepresented in the levels of

informal experiences' engagement. We will check to which extent students' research interest differs between these three groups.

Finally, since our research question is aimed at 1st-year students, we assess scientific thinking on a level that might be seen as quite basic for a higher education level. Measures on a more advanced level could be added, for example assessing justification in multiple sources as another facet of epistemic cognition (Ferguson and Bråten, 2013). More advanced measures might add relevant information, but they might potentially also mostly reveal floor effects due to 1st-year students' limited experiences.

ETHICS STATEMENT

The study materials and consent forms have been developed in accordance with ethical norms and guidelines from all participating countries and universities. The study was approved by the Institutional Review Board of Istanbul Bilgi University and IADT Institute Research Ethics Committee.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENT

This project is conducted under Junior Researcher Programme.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2017.00133/full#supplementary-material>

REFERENCES

- Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: Freeman.
- Barron, B. (2006). Interest and self-sustained learning as catalysts of development: a learning ecology perspective. *Hum. Dev.* 49, 193–224. doi: 10.1159/000094368
- Beißert, H., Köhler, M., Rempel, M., and Beierlein, C. (2014). *Eine Deutschsprachige Kurzsкала zur Messung des Konstrukts Need for Cognition: Die Need for Cognition Kurzsкала (NFC-K). GESIS-Working Papers 2014/32*. Available at: <http://nbn-resolving.de/urn:nbn:de:0168-ssolar-403157>
- Bråten, I., and Ferguson, L. E. (2014). Investigating cognitive capacity, personality, and epistemic beliefs in relation to science achievement. *Learn. Individ. Differ.* 36, 124–130. doi: 10.1016/j.lindif.2014.10.003
- Bråten, I., and Strømso, H. I. (2005). The relationship between epistemological beliefs, implicit theories of intelligence, and self-regulated learning among norwegian postsecondary students. *Br. J. Educ. Psychol.* 75, 539–565. doi: 10.1348/000709905X25067
- Cacioppo, J. T., and Petty, R. E. (1982). The need for cognition. *J. Pers. Soc. Psychol.* 42, 116–131. doi: 10.1037/0022-3514.42.1.116
- Carini, R. M., Kuh, G. D., and Klein, S. P. (2006). Student engagement and student learning: testing the linkages. *Res. High. Educ.* 47, 1–32. doi: 10.1007/s11162-005-8150-9
- Caro, D., and Cortes, D. (2012). Measuring family socioeconomic status: an illustration using data from PIRLS 2006. *IERI Monogr. Ser.* 5, 9–33.
- Chemers, M. M., Hu, L.-T., and Garcia, B. F. (2001). Academic self-efficacy and first-year college student performance and adjustment. *J. Educ. Psychol.* 93, 55–64. doi: 10.1037/0022-0663.93.1.55
- Chen, J. A., Metcalf, S. J., and Tutwiler, M. S. (2014). Motivation and beliefs about the nature of scientific knowledge within an immersive virtual ecosystems environment. *Contemp. Educ. Psychol.* 39, 112–123. doi: 10.1016/j.cedpsych.2014.02.004
- Chen, J. A., and Usher, E. L. (2013). Profiles of the sources of science self-efficacy. *Learn. Individ. Differ.* 24, 11–12. doi: 10.1016/j.lindif.2012.11.002
- Coleman, A. B., Lam, D. P., and Soowal, L. N. (2015). Correlation, necessity, and sufficiency: common errors in the scientific reasoning of undergraduate students for interpreting experiments. *Biochem. Mol. Biol. Educ.* 43, 305–315. doi: 10.1002/bmb.20879
- Dabbagh, N., and Kitsantas, A. (2012). Personal learning environments, social media, and self-regulated learning: a natural formula for connecting formal

- and informal learning. *Internet High. Educ.* 15, 3–8. doi: 10.1016/j.iheduc.2011.06.002
- Drummond, C., and Fischhoff, B. (2015). Development and validation of the scientific reasoning scale. *J. Behav. Decis. Mak.* 30, 26–38. doi: 10.1002/bdm.1906
- Duffy, S. (2010). Random numbers demonstrate the frequency of Type I errors: three spreadsheets for class instruction. *J. Stat. Educ.* 18, 1–16.
- Duschl, R. A., and Grandy, R. (2013). Two views about explicitly teaching nature of science. *Sci. Educ.* 22, 2109–2139. doi: 10.1007/s11191-012-9539-4
- Ebner, M., Lienhardt, C., Rohs, M., and Meyer, I. (2010). Microblogs in higher education - a chance to facilitate informal and process-oriented learning? *Comput. Educ.* 55, 92–100. doi: 10.1016/j.compedu.2009.12.006
- Elias, S. M., and Loomis, R. J. (2002). Utilizing need for cognition and perceived self-efficacy to predict academic performance. *J. Appl. Soc. Psychol.* 32, 1687–1702. doi: 10.1111/j.1559-1816.2002.tb02770.x
- Eshach, H. (2007). Bridging in-school and out-of-school learning: formal, non-formal, and informal education. *J. Sci. Educ. Technol.* 16, 171–190. doi: 10.1007/s10956-006-9027-1
- Etz, A., Gronau, Q. F., Dablander, F., Edelsbrunner, P. A., and Baribault, B. (2016). *How to Become a Bayesian in Eight Easy Steps: An Annotated Reading List*. Available at: <http://tinyurl.com/eightstepsV1>
- Evans, M. D. R., Kelley, J., Sikora, J., and Treiman, D. J. (2010). Family scholarly culture and educational success: books and schooling in 27 nations. *Res. Soc. Stratif. Mobil.* 28, 171–197. doi: 10.1016/j.rssm.2010.01.002
- Falk, J. H. (2001). *Free-Choice Science Education: How We Learn Science outside of School. Ways of Knowing in Science and Mathematics Series*. Available at: <https://eric.ed.gov/?id=ED465540>
- Ferguson, L. E., and Bråten, I. (2013). Student profiles of knowledge and epistemic beliefs: changes and relations to multiple-text comprehension. *Learn. Instr.* 25, 49–61. doi: 10.1016/j.learninstruc.2012.11.003
- Franco, G. M., Muis, K. R., Kendeou, P., Ranellucci, J., Sampasivam, L., and Wang, X. (2012). Examining the influences of epistemic beliefs and knowledge representations on cognitive processing and conceptual change when learning physics. *Learn. Instr.* 22, 62–77. doi: 10.1016/j.learninstruc.2011.06.003
- Gelman, A., and Stern, H. (2006). The difference between ‘significant’ and ‘not significant’ is not itself statistically significant. *Am. Stat.* 60, 328–331. doi: 10.1198/000313006X152649
- Gigerenzer, G. (2004). Mindless statistics. *J. Socio Econ.* 33, 587–606. doi: 10.1016/j.jsocec.2004.09.033
- Gokhale, A. A. (1995). Collaborative learning enhances critical thinking. *J. Technol. Educ.* 7, 22–30. doi: 10.21061/jte.v7i1.a.2
- Greene, J. A., Azevedo, R., and Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: philosophical perspectives and methodological directions. *Educ. Psychol.* 43, 142–160. doi: 10.1080/00461520802178458
- Greene, J. A., Torney-Purta, J., and Azevedo, R. (2010). Empirical evidence regarding relations among a model of epistemic and ontological cognition, academic performance, and educational level. *J. Educ. Psychol.* 102, 234–255. doi: 10.1037/a0017998
- Greene, J. A., and Yu, S. B. (2014). Modeling and measuring epistemic cognition: a qualitative re-investigation. *Contemp. Educ. Psychol.* 39, 12–28. doi: 10.1016/j.cedpsych.2013.10.002
- Halpern, D. F., Millis, K., Graesser, A. C., Butler, H., Forsyth, C., and Cai, Z. (2012). Operation ARA: a computerized learning game that teaches critical thinking and scientific reasoning. *Think. Skills Creat.* 7, 93–100. doi: 10.1016/j.tsc.2012.03.006
- Haslbeck, J. M. B., and Waldorp, L. J. (2016). *Mgm: Structure Estimation for Time-Varying Mixed Graphical Models in High-Dimensional Data*. Available at: <http://arxiv.org/abs/1510.06871>
- Hoekstra, R., Morey, R. D., Rouder, J. N., and Wagenmakers, E.-J. (2014). Robust misinterpretation of confidence intervals. *Psychon. Bull. Rev.* 21, 1157–1164. doi: 10.3758/s13423-013-0572-3
- Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemp. Educ. Psychol.* 25, 378–405. doi: 10.1006/ceps.1999.1026
- Hofer, B. K., and Pintrich, P. R. (1997). The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Rev. Educ. Res.* 67, 88–140. doi: 10.3102/00346543067001088
- Hofstein, A., and Rosenfeld, S. (1996). Bridging the gap between formal and informal science learning. *Stud. Sci. Educ.* 28, 87–112. doi: 10.1080/03057269608560085
- Holmes, N. G., Wieman, C. E., and Bonn, D. A. (2015). Teaching critical thinking. *Proc. Natl. Acad. Sci. U.S.A.* 112, 11199–11204. doi: 10.1073/pnas.1505329112
- Jarosoz, A. F., and Wiley, J. (2014). What are the odds? A practical guide to computing and reporting bayes factors. *J. Probl. Solving* 7:2. doi: 10.7771/1932-6246.1167
- Jenert, T., Postareff, L., Brahm, T., and Sari, L.-Y. (2015). Editorial: enculturation and development of beginning students. *Zeitschrift Für Hochschulentwicklung* 10, 9–21.
- Johnson, D. W., Johnson, R. T., and Smith, K. A. (1998). Cooperative learning returns to college what evidence is there that it works? *Change* 30, 26–35. doi: 10.1080/00091389809602629
- Kassens-Noor, E. (2012). Twitter as a teaching practice to enhance active and informal learning in higher education: the case of sustainable tweets. *Active Learn. High. Educ.* 13, 9–21. doi: 10.1177/1469787411429190
- Kitchener, R. F. (2002). Folk epistemology: an introduction. *New ideas Psychol.* 20, 89–105. doi: 10.1016/S0732-118X(02)00003-X
- Klahr, D. (2000). *Exploring science: The Cognition and Development of Discovery Processes*. Cambridge, MA: MIT press.
- Klahr, D., Zimmerman, C., and Jirout, J. (2011). Educational interventions to advance children’s scientific thinking. *Science* 333, 971–975. doi: 10.1126/science.1204528
- Krause, K.-L., and Coates, H. (2008). Students’ engagement in first-year university. *Assess. Eval. High. Educ.* 22, 493–505. doi: 10.1080/02602930701698892
- Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., and Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *J. High. Educ.* 79, 540–563. doi: 10.1353/jhe.0.0019
- Kuhn, D. (2002). “What Is Scientific Thinking and How Does It Develop?,” in *Blackwell Handbook of Childhood Cognitive Development*, ed. U. Goswami (Hoboken, NJ: Blackwell Publishers Ltd), 371–393.
- Kuhn, D. (2009). Do students need to be taught how to reason? *Educ. Res. Rev.* 4, 1–6. doi: 10.1016/j.edurev.2008.11.001
- Kuhn, D., Cheney, R., and Weinstock, M. (2000). The development of epistemological understanding. *Cogn. Dev.* 15, 309–328. doi: 10.1016/S0885-2014(00)00030-7
- Kuhn, D., Ramsey, S., and Arvidsson, T. S. (2015). Developing multivariable thinkers. *Cogn. Dev.* 35, 92–110. doi: 10.1016/j.cogdev.2014.11.003
- Laurillard, D. (2009). The pedagogical challenges to collaborative technologies. *Int. J. Comput. Support. Collab. Learn.* 4, 5–20. doi: 10.1007/s11412-008-9056-2
- Lawson, A. E., Clark, B., Cramer-Meldrum, E., Falconer, K. A., Sequist, J. M., and Kwon, Y.-J. (2000). Development of scientific reasoning in college biology: Do two levels of general hypothesis-testing skills exist? *J. Res. Sci. Teach.* 37, 81–101.
- Lehman, D. R., and Nisbett, R. E. (1990). A longitudinal study of the effects of undergraduate training on reasoning. *Dev. Psychol.* 26, 952–960. doi: 10.1037/0012-1649.26.6.952
- Manning, K., Zachar, P., Ray, G., and LoBello, S. (2006). Research methods courses and the scientist and practitioner interests of psychology majors. *Teach. Psychol.* 33, 194–196.
- Mason, L. (2016). “Psychological perspectives on measuring epistemic cognition,” in *Handbook of Epistemic Cognition*, eds J. A. Greene, W. A. Sandoval, and I. Bråten (New York, NY: Routledge), 375–392. doi: 10.4324/9781315795225
- McElreath, R. (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Boca Raton, FL: CRC Press.
- Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., and Wagenmakers, E.-J. (2015). The fallacy of placing confidence in confidence intervals. *Psychon. Bull. Rev.* 23, 103–123. doi: 10.3758/s13423-015-0947-8
- Morey, R. D., and Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychol. Methods* 16, 406–419. doi: 10.1037/a0024377
- Moss, E. (2012). *Assessing Understanding of the Nature of Science and Science Self-Efficacy in Undergraduates Involved in Research in an Introductory Geology Course*. Available at: <http://lib.dr.iastate.edu/etd/12825>
- Muis, K. R., and Franco, G. M. (2009). Epistemic profiles and metacognition: support for the consistency hypothesis. *Metacogn. Learn.* 5, 27–45. doi: 10.1007/s11409-009-9041-9

- Nieuwenhuis, S., Forstmann, B. U., and Wagenmakers, E.-J. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nat. Neurosci.* 14, 1105–1107. doi: 10.1038/nn.2886
- Nussbaum, E. M., Sinatra, G. M., and Poliquin, A. (2008). Role of epistemic beliefs and scientific argumentation in science learning. *Int. J. Sci. Educ.* 30, 1977–1999. doi: 10.1080/09500690701545919
- Parent, M. C., and Oliver, J. A. (2015). Mentoring the earliest-career psychologists: role models, knowledge of internship issues, and attitudes toward research and science. *Prof. Psychol. Res. Pr.* 46, 55–61. doi: 10.1037/a0038839
- Pashler, H., and Wagenmakers, E.-J. (2012). Editors' introduction to the special section on replicability in psychological science a crisis of confidence? *Perspect. Psychol. Sci.* 7, 528–530. doi: 10.1177/1745691612465253
- Patrick, W. (2010). *Recognising Non-Formal and Informal Learning Outcomes, Policies and Practices: Outcomes, Policies and Practices*. Paris: OECD Publishing.
- Plummer, M. (2003). "JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling," in *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, Vienna, Vol. 124, 125.
- Plummer, M. (2013). *rjags: Bayesian Graphical Models Using MCMC. R Package Version 3*.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 409.
- Resnick, L. B. (1987). The 1987 presidential address learning in school and out. *Educ. Res.* 16, 13–54. doi: 10.3102/0013189X016009013
- Riggs, M. L., Warka, J., Babasa, B., Betancourt, R., and Hooker, S. (1994). Development and validation of self-efficacy and outcome expectancy scales for job-related applications. *Educ. Psychol. Meas.* 54, 793–802. doi: 10.1177/0013164494054003026
- Schönbrodt, F. D., and Wagenmakers, E.-J. (2016). *Bayes Factor Design Analysis: Planning for Compelling Evidence*. SSRN Scholarly Paper ID 2722435. Rochester, NY: Social Science Research Network.
- Sodian, B., Zaitchik, D., and Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Dev.* 62, 753–766. doi: 10.1111/j.1467-8624.1991.tb01567.x
- Stoloff, M., McCarthy, M., Keller, L., Varfolomeeva, V., Lynch, J., Makara, K., et al. (2009). The undergraduate psychology major: an examination of structure and sequence. *Teach. Psychol.* 37, 4–15. doi: 10.1080/00986280903426274
- Su, Y.-S., and Yajima, M. (2012). *R2jags: A Package for Running jags from R. R Package Version 0.03-08*. Available at: <http://CRAN.R-project.org/package=R2jags>
- Van De Schoot, R., Kluytmans, A., Tummers, L., Lugtig, P., Hox, J., and Muthén, B. (2012). Facing off with scylla and charybdis: a comparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Front. Psychol.* 4:770. doi: 10.3389/fpsyg.2013.00770
- VanderStoep, S. W., and Shaughnessy, J. J. (1997). Taking a course in research methods improves reasoning about real-life events. *Teach. Psychol.* 24, 122–124. doi: 10.1207/s15328023top2402_8
- VanWormer, L. A., Jordan, E. F., and Blalock, L. D. (2014). Assessing the perceived value of research participation. *Teach. Psychol.* 41, 233–236. doi: 10.1177/0098628314537974
- Vivian, R., Barnes, A., Geer, R., and Wood, D. (2014). The academic journey of university students on facebook: an analysis of informal academic-related activity over a semester. *Res. Learn. Technol.* 22:24681. doi: 10.3402/rlt.v22.24681
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychon. Bull. Rev.* 14, 779–804. doi: 10.3758/BF03194105
- Wagenmakers, E.-J., Morey, R. D., and Lee, M. D. (2016). Bayesian benefits for the pragmatic researcher. *Curr. Dir. Psychol. Sci.* 25, 169–176. doi: 10.1177/0963721416643289
- Walker, C. O., Greene, B. A., and Mansell, R. A. (2006). Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement. *Learn. Individ. Differ.* 16, 1–12. doi: 10.1016/j.lindif.2005.06.004
- Wayment, H. A., and Dickson, K. L. (2008). Increasing student participation in undergraduate research benefits students, faculty, and department. *Teach. Psychol.* 35, 194–197. doi: 10.1177/009862830803500307
- Woo, S. E., Harms, P. D., and Kuncel, N. R. (2007). Integrating personality and intelligence: typical intellectual engagement and need for cognition. *Pers. Individ. Dif.* 6, 1635–1639. doi: 10.1016/j.paid.2007.04.022
- Yang, M. (2015). *Promoting Self-Sustained Learning in Higher Education: The ISEE Framework*. Available at: <http://repository.lib.eduhk.hk/jspui/handle/2260/17647>
- Ziegler, M., Kemper, C. J., and Lenzner, T. (2015). The issue of fuzzy concepts in test construction and possible remedies. *Eur. J. Psychol. Assess.* 31, 1–4. doi: 10.1027/1015-5759/a000255
- Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Dev. Rev.* 27, 172–223. doi: 10.1016/j.dr.2006.12.001

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Soyylmaz, Griffin, Martín, Kucharský, Peycheva, Vaupotič and Edelsbrunner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.