

# Beyond Structure: Using the Rational Force Model to Assess Argumentative Writing

Written Communication

1–31

© 2023 SAGE Publications



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/07410883221148664

[journals.sagepub.com/home/wcx](https://journals.sagepub.com/home/wcx)

Ylva Backman<sup>1</sup>, Alina Reznitskaya<sup>2</sup>,  
Viktor Gardelli<sup>1</sup>, and Ian A. G. Wilkinson<sup>3</sup>

## Abstract

Current approaches used in educational research and practice to evaluate the quality of written arguments often rely on structural analysis. In such assessments, credit is awarded for the presence of structural elements of an argument, such as claims, evidence, and rebuttals. In this article, we discuss limitations of such approaches, including the absence of criteria for evaluating the quality of the argument elements. We then present an alternative framework, based on the Rational Force Model (RFM), which originated from the work of a Nordic philosopher Næss. Using an example of an argumentative essay, we demonstrate the potential of the RFM to improve argument analysis by focusing on the acceptability and relevance of argument elements, two criteria widely considered to be fundamental markers of argument strength. We outline possibilities and challenges with using the RFM in educational contexts and conclude by proposing directions for future research.

## Keywords

argument analysis, argumentation, assessment framework, informal logic, writing assessment

<sup>1</sup>Department of Health, Education and Technology, Luleå University of Technology, Sweden

<sup>2</sup>Department of Educational Foundations, Montclair State University, NJ, USA

<sup>3</sup>Department of Teaching and Learning, Ohio State University, Columbus, USA

## Corresponding Author:

Viktor Gardelli, Department of Health, Education and Technology, Luleå University of Technology, Luleå, 97187, Sweden.

Email: [viktor.gardelli@ltu.se](mailto:viktor.gardelli@ltu.se)

*[Schools need] to cultivate deep-seated and effective habits of discriminating tested beliefs from mere assertions, guesses, and opinions; to develop a lively, sincere, and open-minded preference for conclusions that are properly grounded, and to ingrain into the individual's working habits methods of inquiry and reasoning appropriate to the various problems that present themselves.*

Dewey (1910, p. 28)

In academic circles and beyond, people are becoming increasingly concerned about the phenomenon of “post-truth” and its eroding effects on the functioning of liberal democracies (Barzilai & Chinn, 2020; Crosswhite, 2018; Kavanagh & Rich, 2018). Briefly, post-truth is a societal condition characterized by an increased volume of falsehoods, a declining trust in information provided by established institutions, and a growing disagreement about verifiable facts (Kavanagh & Rich, 2018). A vivid example of the serious consequences of post-truth for society and individual well-being is a recent struggle to coordinate an evidence-based response to the COVID-19 pandemic. Despite ongoing efforts of the scientific community to share disciplinary knowledge about infections, treatments, and vaccinations, many people around the world continue to resist vaccination efforts based on unsubstantiated speculations and conspiracy theories about the disease (Bolsen & Palm, 2022; Puri et al., 2020).

Education has long been seen as a powerful mechanism to combat misinformation and prepare future citizens to make better decisions (Dewey, 1910; Postman, 1995). Reflecting this view, national and international educational policy documents have identified competency in argumentation as an important educational goal and stressed the importance of providing students with opportunities to formulate and justify their own views and to critically evaluate those of others (National Governors Association, 2010; Partnership for 21st Century Skills, 2012; Swedish National Agency for Education, 2021). There is also a growing body of theory and research on instructional interventions that can promote student engagement in argumentation in a variety of disciplines (Asterhan & Schwarz, 2016; Kuhn, 2010; Osborne, 2010; Reznitskaya et al., 2009).

Yet even our best efforts to date may not be sufficiently rigorous to help students learn how to navigate the increasingly disorienting informational landscape created by the post-truth condition. In an insightful article on this topic, Barzilai and Chinn (2020) proposed that current education might, in fact, contribute to the development of post-truth thinking by not focusing enough on the epistemic criteria that underlie sound judgments. That is, although students may now be more encouraged to share their opinions, less

attention is being paid to discussing how we know what we know and what makes some conclusions more reasonable than others. We further suggest that practices that promote the sharing of opinions without critically examining them may reinforce poor reasoning and contribute to the development of relativist epistemic beliefs, according to which knowledge is seen as entirely subjective and idiosyncratic and there are no established criteria that can be used to decide on the trustworthiness of knowledge claims.

Moreover, typical assessment practices used to evaluate students' performance on argument-related tasks, such as an argumentative essay, often rely on structural analysis. That is, they omit important criteria of argument evaluation, focusing instead on the presence—rather than quality—of structural elements, such as claims, evidence, and rebuttals (Chinn et al., 2016; Newell et al., 2011; Rapanta et al., 2013). Relying on such assessments may prevent teachers and students from acquiring deep understanding of how to distinguish between good and poor reasoning. Thus, current assessment practices might further contribute to the development of relativist conceptions about knowledge construction and justification.

The purpose of this article is to explore the use of a different analytic framework, which goes beyond structural analysis. The framework, called the Rational Force Model (RFM), originated from the work of a Norwegian philosopher Næss (1959) and was further developed by several Nordic scholars (Backman et al., 2012; Björnsson et al., 1994; Føllesdal et al., 1990; Gardelli et al., 2019). We and others (Backman et al., 2012; Björnsson et al., 2009) have refined the RFM into the commonly accepted version described here and used it to teach college-level philosophy courses. In this article, we introduce the RFM to a broader audience and explore the ways in which it can inform new approaches to assessment of written arguments in educational contexts.

We begin with a review of common practices used to evaluate argument quality in educational contexts, focusing on the limitations of assessments based largely on structural analysis. Next, we present principles and procedures of the RFM. Using an example of an argumentative essay, we demonstrate the potential of the RFM to improve argument analysis and contribute to the diagnostic and instructional value of educational assessments. We then outline the possibilities and challenges related to the use of the RFM in educational settings and conclude by proposing directions for future research.

## **Limitations of Structural Analysis for Assessing Argument Quality**

Structural analysis is commonly used in educational research and practice to evaluate the quality of students' arguments (Erduran et al., 2004; Rapanta

et al., 2013; Sampson & Clark, 2008). In such analysis, many educators rely on Toulmin's (1958) model, which describes the general structure of an argument and identifies unique functions of six core elements: claims, grounds, warrants, backings, qualifiers, and rebuttals. When assessing arguments using Toulmin's model, credit is typically awarded for the presence of select argument elements. For example, McNeill and colleagues (McNeill, 2011; McNeill & Knight, 2013; McNeill & Krajcik, 2012) drew on Toulmin's model to develop a CER framework with three argument elements: Claims, Evidence, and Reasoning. This framework informed assessment, instruction, and teacher professional development aimed at supporting scientific argumentation in middle school science classrooms.

Structural analysis, either based on Toulmin's model or other frameworks (e.g., Kuhn & Crowell, 2011), offers a useful heuristic for assessing student arguments. Indeed, studies have shown that structural analysis can account for a large proportion of variance in argument quality (De La Paz et al., 2012; Ferretti et al., 2009). However, several researchers have criticized the educational community's reliance on this approach, pointing out practical and conceptual limitations (Chinn et al., 2016; Erduran, 2018; Kim & Roth, 2018; Nussbaum et al., 2019; Oyler, 2019). Assessments based on structural analysis, including those used in our own studies, typically reward students for producing more of the desired argument elements in their essays (Dong et al., 2008; Erduran et al., 2004; Reznitskaya et al., 2001; Zohar & Nemet, 2002). Although such assessments identify key elements of an argument, they do not provide criteria for judging the *quality* of these elements (Nussbaum et al., 2019). Thus, even if structural analysis captures variability in student responses, it does not focus on epistemic criteria that underlie sound judgments, and therefore has limited diagnostic value for informing assessment, feedback, and instruction.

In a compelling example, Chinn et al. (2016) illustrated problems with relying solely on a structural analysis to assess the quality of arguments. Using the written argument against vaccinations shown below, they demonstrated how seriously flawed reasoning can remain undetected, and even receive high scores based on the presence of key structural elements:

Children should not be vaccinated. A study published by Wakefield and colleagues in *British Medical Journal*, one of the world's most prestigious journals, showed that children given MMR vaccinations were much more likely to become autistic. This means that vaccinations are dangerous and cause autism. Doctors say that my children will be safer if given a vaccination, but they do not have personal knowledge of my children. I do, and I know in my heart that they are better off without taking that terrible risk. Scientists say that

we should trust their studies that show that vaccinations do not cause autism, but the scientists who conducted them expected the vaccinations to work, so they are quite biased and should not be taken seriously. Now lots of US television programs are shedding light on this issue, which shows that there are more and more experts who are coming around and opposing vaccinations and debating the scientific establishment. This shows that there is now a strong expert opposition to vaccination.

As Chinn et al. (2016) noted, this text “scores at the pinnacle” (p. 8) of any structural analysis. It incorporates several desirable argument elements, including claims (e.g., *Children should not be vaccinated*), grounds (e.g., *A study. . . showed that children given MMR vaccinations were much more likely to become autistic*), and warrants (e.g., *This means that vaccinations are dangerous and cause autism*). The text also contains counterarguments (e.g., *Doctors say that my children will be safer if given a vaccination*), which are undermined by the use of rebuttals (e.g., . . . *but they do not have personal knowledge of my children*). However, the quality of all these structural elements and their contribution to the overall strength of the argument are not typically scrutinized in a structural analysis, allowing inaccurate and irrelevant justifications to go unnoticed.

Some researchers have combined structural analysis with other schemes designed to take into account the content of arguments, thus enhancing the validity of assigned scores (e.g., Kuhn & Udell, 2003; Means & Voss, 1996; Reznitskaya & Wilkinson, 2020). For example, when assessing student arguments about everyday ill-structured problems, Means and Voss (1996) supplemented structural analysis with a scheme that categorized reasons into six categories: abstract, consequential, rule-based, authority, personal, and vague. Although the authors assumed that the quality of reasons decreased in the order of the presented categories, they did not offer any justification for the proposed hierarchy. In our own work, when assessing persuasive essays of elementary school students, we expanded our structural analysis by incorporating a list of acceptable and relevant statements for and against the main claim (Reznitskaya & Wilkinson, 2020). We developed this list by first conducting a thematic analysis of a sample of students’ essays, and we awarded credit to statements in students’ essays only if they appeared on the list. Such additions to purely structural analysis are helpful, but they are atheoretical and inefficient (i.e., they are content specific). Moreover, they fail to account for the differential quality of argument elements used by students to strengthen or undermine the positions in their arguments. Also, because the elements are counted in a piecemeal fashion, the relationships among elements remain unexamined.

One notable alternative to structural analysis in educational contexts is the classification of argument schemes described by Walton (Walton, 1996a; Walton et al., 2008). It includes descriptions of more than 60 reasoning patterns commonly used in a variety of everyday and disciplinary contexts. Each argument scheme is accompanied by a list of critical questions that can be asked to evaluate arguments following the scheme in a given context. For example, when encountering a scheme called an “argument from expert opinion,” one can ask whether the expert is credible, whether the knowledge of the expert is in the relevant domain, and whether the expert’s claim is supported by evidence (Walton, 2014). Asking such critical questions can inform the analysis of several propositions in the anti-vaccination text, such as *Doctors . . . do not have personal knowledge of my children. I do, and I know in my heart that they are better off without taking that terrible risk.* In these propositions, the author is contrasting the knowledge of a doctor with that of a parent. Raising critical questions about the relevant expertise and the kind of evidence most appropriate to determine the safety of medical procedures can help identify weaknesses in reasoning that values parental knowledge over medical expertise.

Walton’s argument schemes offer helpful and theoretically grounded analytic tools, and educational researchers are increasingly using them to design innovative instructional and assessment applications in a variety of school subjects, including science, history, social studies, and math (for a review, see Rapanta, 2022). For example, Ferretti et al. (2009) examined the use of argument schemes in the persuasive essays of elementary school children. The authors applied normative standards for specific argument schemes that were appropriate for the given discursive purpose (i.e., to persuade) and controversial topic (i.e., deciding on a school policy). They suggested that students should use Walton’s (1996b) “argument from consequence” scheme and listed relevant critical questions that students needed to consider in their essays, such as “How sure are you that the (good, bad) consequences (outcomes, results) will actually happen? How do you know that these consequences will actually happen? Do you have evidence (facts, data, support) that these consequences probably will happen if we implement the policy?” (Ferretti et al., 2009, p. 587). Such applications of Walton’s schemes help to address the key problem of structural analysis, as they refocus educators’ efforts on the critical evaluation of argument elements, rather than their mere presence. Still, Walton’s critical questions do not provide a direct assessment of argument quality, as there are no established procedures for converting critical questions into judgments regarding the reasonableness of an argument.

In sum, although assessments based on structural analysis offer practical tools that capture certain differences in student performance, they provide

limited diagnostic and instructional value in terms of the epistemic criteria for evaluating the quality of an argument. Alternative evaluation schemes are either ad hoc or, if theoretically grounded, do not readily make visible the basis for judging argument quality. In the next section, we introduce another promising analytic framework for evaluating arguments, the RFM, which enables a theoretically sound, systematic, and transparent analysis of arguments based on key epistemic criteria of acceptability and relevance. We suggest it could provide a valuable framework to improve assessment and inform instruction in various educational contexts.

## The Rational Force Model

### *Overview of the Analytic Process*

An RFM analysis relies on two key concepts: acceptability and relevance. The *acceptability* of a proposition reflects the degree to which we have good reasons to believe that the proposition is correct. The *relevance* of a proposition reflects the degree to which we have good reasons to believe that the proposition, if it were correct, would be effective in achieving its aim.

The RFM analysis of an argument is conducted in two phases: *Descriptive* and *Evaluative*. The Descriptive phase precedes the Evaluative phase and entails two interpretive processes: a *propositional reconstruction*, where the wording from the original text is reformulated into grammatically correct sentences that express distinct ideas, and an *argument reconstruction*, where the relations between these ideas are analyzed and shown using a specific notation.

During *argument reconstruction*, we differentiate between first-level propositions, directed towards the main claim, and other propositions (second-, third-level, etc.), which are directed toward propositions other than the main claim. Each proposition is then classified as intended to either support or undermine the acceptability or relevance of its target proposition. Thus, an RFM argument reconstruction is based on a classification with four types of propositions in addition to the main claim, as shown in Figure 1.

As shown in Figure 1, the classification used during the argument reconstruction is based on an interpretation of the aims of the propositions. For instance, an acceptability support is defined as a proposition intended to increase another proposition's acceptability, and the classification in the Descriptive phase makes no promise as to whether a proposition *actually* succeeds in supporting another proposition's acceptability.

In the second, Evaluative phase, each proposition resulting from the Descriptive phase (except for the main claim) is evaluated using the same

	Support	Objection
Acceptability	<p><i>Acceptability support (AS):</i> A proposition intended to increase another proposition's acceptability</p>	<p><i>Acceptability objection (AO):</i> A proposition intended to decrease another proposition's acceptability</p>
Relevance	<p><i>Relevance support (RS):</i> A proposition intended to increase another proposition's relevance</p>	<p><i>Relevance objection (RO):</i> A proposition intended to decrease another proposition's relevance</p>

**Figure 1.** Four types of propositions in an RFM analysis.

Note. Adapted from Backman et al. (2012).

concepts of acceptability and relevance that were used to classify the propositions during the argument reconstruction. Importantly, the relevance of a proposition during the evaluation process is assessed independently of its acceptability. For instance, if one argued that “School is dangerous for most students” by stating “All teachers are zombies,” we would assign a very low acceptability value to the latter proposition, because we have no good reasons to believe that it is correct. Nonetheless, we would assign a very high relevance value to the very same proposition, since *if it were* correct that all teachers were zombies, then we would have good reasons for believing that school is dangerous.

During the Evaluative phase, each proposition, regardless of its type, is assigned three values: an *acceptability* (A) value, a *relevance* (R) value, and, derived from these, a *rational force* (RF) value. In various versions of the RFM, scholars have relied on different scales and analytic schemes for assigning these values (Backman et al., 2012; Björnsson et al., 1994, 2009). To make the RFM more accessible, it is common to first present the model using an ordinal scale. In this article, following Björnsson et al. (2009), we use a 5-point scale, with values ranging from *very low* (VL), indicating that the value is close to nonexistent, to *very high* (VH), indicating that the value is as close to maximal as possible.

The *rational force* value is the strength with which each proposition in the argument influences the proposition it is aimed at, and it is fully determined by the acceptability and relevance values. For instance, an acceptability support with a very high rational force positively affects the acceptability of the proposition that it is intended to affect. Likewise, a relevance support with a very high rational force positively affects the relevance of the proposition that it is intended to affect. Thus, both acceptability and relevance values are influenced by the rational force values of higher-level propositions. When all propositions except for the main claim have been evaluated and assigned a



Prop. No.	Argument reconstruction	Verbatim	Propositional reconstruction	Evaluation		
				A	R	RF
1	T	Children should not be vaccinated.		VL		
2	AS <sub>1</sub>	"This means that vaccinations are dangerous ..."	Vaccinations are dangerous.	VL	H	VL
3	AS <sub>1</sub> AS <sub>1</sub>	"... and cause autism."	Vaccinations cause autism.	VL	H	VL
4	AS <sub>1</sub> AS <sub>1</sub> AS <sub>1</sub>	"A study published by Wakefield and colleagues in British Medical Journal ... showed that children given MMR vaccinations were much more likely to become autistic."	A study in the BMJ shows that children given MMR vaccinations were much more likely to become autistic.	VL	H	VL
5	RS <sub>1</sub> AS <sub>1</sub> AS <sub>1</sub> AS <sub>1</sub>	"... one of the world's most prestigious journals ..."	BMJ is one of the world's most prestigious journals.	VH	H	H

**Figure 2.** RFM analysis of an excerpt from the anti-vaccination text.

Note. "R" stands for relevance, "A" for acceptability, and "RF" for rational force. We use a scale with the following values for relevance, acceptability, and rational force: very low (VL), low (L), medium (M), high (H), and very high (VH).

rational force value, it is possible to ultimately weigh together the rational forces of all first-level propositions and assign a single *acceptability* value to the main claim. This acceptability value represents the cumulative rational force of all other propositions in an argument.

The results from both the Descriptive and the Evaluative phases of the RFM analysis are simultaneously presented in a table. To illustrate the RFM analysis, let us consider the first few lines from the anti-vaccination text, shown in Figure 2 (analysis of the entire text is presented in Appendix). In row 1, we present the main claim (MC) that "children should not be vaccinated." In row 2, in the "Verbatim" column, we show the original wording of a proposition supporting the main claim ("This means that vaccinations are dangerous . . ."). To the right, in the column called "Propositional reconstruction," we capture the intended meaning of the original wording, using a full sentence ("Vaccinations are dangerous"). In the "Argument reconstruction" column, row 2, "AS<sub>1</sub>" denotes that this is the 1st *acceptability support* for the main claim. The notation "AS<sub>1</sub>AS<sub>1</sub>" in the next row denotes the 1st *acceptability support* for AS<sub>1</sub>. That is, the proposition "Vaccinations cause autism" is a second-level proposition aimed at making the first-level proposition "Vaccinations are dangerous" more acceptable. The notation "RS<sub>1</sub>AS<sub>1</sub>AS<sub>1</sub>AS<sub>1</sub>" in row 5 marks the first *relevance support*. That is, the statement "BMJ is one of the world's most prestigious journals" is intended to enhance the relevance of the results of a study reported in this journal (AS<sub>1</sub>AS<sub>1</sub>AS<sub>1</sub>) with regard to the claim that "Vaccinations cause autism" (AS<sub>1</sub>AS<sub>1</sub>).

In the "Evaluation" columns to the right, we present results from the Evaluative phase, where we assigned three values to each proposition: an

*acceptability* (A) value, a *relevance* (R) value, and, derived from these, a *rational force* (RF) value. To evaluate the acceptability and relevance of a proposition, we must know its relation to other propositions in the argument, which is why the Descriptive phase precedes the Evaluative phase. As noted earlier, both acceptability and relevance values are influenced by the rational force values of higher-level propositions. For example, the acceptability of the lower-level proposition  $AS_1$  (“vaccinations are dangerous”) is affected by the rational force scores of higher-level propositions directed towards them, such as  $AS_1AS_1$  (“vaccinations cause autism”). Note that, although  $AS_1$  is highly relevant to the main claim, it has a very low acceptability because of the low rational force value of its supporting proposition,  $AS_1AS_1$ . The very low acceptability value of  $AS_1$  indicates that there are no compelling reasons to trust the statement that vaccinations are dangerous.

In the final part of the Evaluative phase, the main claim is assigned a single acceptability value, which reflects the aggregated rational force scores of all previous propositions. In Figure 2, the final result of the Evaluative phase is a very low acceptability value of the main claim (“Children should not be vaccinated”), shown in the first row, under column A.

### Acceptability Values

The *acceptability* value of a proposition is a measure of the degree to which we have reasons to believe that it is correct. The assignment of acceptability values is cumulative. That is, the acceptability of a proposition is affected by the rational forces of acceptability supports and acceptability objections directed at it. For instance, the very low rational force of  $AS_1AS_1AS_1$  (“A study in the *BMJ* shows that children given MMR vaccinations were more likely to become autistic”) reduces the acceptability of  $AS_1AS_1$  (“Vaccinations cause autism”). In turn, the very low rational force value of  $AS_1AS_1$  reduces the acceptability value of  $AS_1$  (“Vaccinations are dangerous”).

Whenever propositions have no acceptability support or acceptability objections directed at them, we must use other strategies for determining their acceptability values. For instance, to assess the *acceptability* of the relevance support “*BMJ* is one of the world’s most prestigious journals” ( $RS_1AS_1AS_1AS_1$  in row 5), we examined databases ranking scientific journals based on impact factors and other bibliometrics. According to these sources, *BMJ* is a highly ranked journal in its subject field, which results in a very high acceptability value.

We also relied on subject-specific knowledge to assess the acceptability of the proposition “A study in the *BMJ* shows that children given MMR vaccinations were much more likely to become autistic” ( $AS_1AS_1AS_1$  in row 4).

According to *BMJ*, the original study reporting a link between vaccinations and autism had serious methodological problems and has been retracted. Based on this knowledge, we assigned a very low acceptability value of  $AS_1AS_1AS_1$ . Moreover, this study was initially published in, and later retracted from, *The Lancet*, not the *BMJ*.

It is important to note that acceptability judgments reflect our current understandings and subjectivities within a given disciplinary context, such as medicine or history. For example, propositions “Heavy smoking causes lung cancer” and “Americans landed on the moon in 1969” should receive very high acceptability values; propositions “Zinc supplements reduce the length of cold” and “Alger Hiss was an intelligence asset of the Soviet government” should receive medium acceptability values; and propositions “Drinking bleach helps cure Covid-19” and “The 9/11 attacks were planned by the US government” should receive very low acceptability values. Björnsson et al. (1994) explained that the result of applying the RFM will “inescapably be influenced by how extensive or limited our knowledge is, and what more or less well-founded valuations and opinions we happen to have. However, this is not a serious problem. The method helps us to determine the validity of an argument in a systematic way. Through that, we can make a judgment about the main claim that is well-founded at least in relation to *our* knowledge, valuations, and opinions.” (p. 51, as translated by the authors). When applying the RFM to educational assessment and instruction, teachers and students can rely on disciplinary expertise and standards to assign acceptability values, and, perhaps more importantly, engage in inquiry about why certain propositions have lower acceptability values than others.

## Relevance Values

The *relevance* value is a measure of the degree to which we have reasons to believe that the proposition, if it were correct, would be effective at achieving its aim. Because effectiveness is dependent on the aim of the proposition, and there are four types of propositions with four different aims, relevance is described differently with regard to each type of proposition. The following definitions are presented in Björnsson et al. (2009, pp. 191-192), with minor changes to terminology (cf. Backman et al., 2012):

- A. The relevance of an acceptability support  $AS$  for another proposition  $P$  is a measure of the extent to which  $AS$  would rationally compel us to accept  $P$ , if we assumed that  $AS$  were correct.
- B. The relevance of an acceptability objection  $AO$  against another proposition  $P$  is a measure of the extent to which  $AO$  would rationally compel us to reject  $P$ , if we assumed that  $AO$  were correct.

- C. The relevance of a relevance support  $RS$  for another proposition  $P$  is a measure of the extent to which  $RS$  would rationally compel us to view  $P$  as relevant, if we assumed that  $RS$  were correct.
- D. The relevance of a relevance objection  $RO$  against another proposition  $P$  is a measure of the extent to which  $RO$  would rationally compel us to view  $P$  as irrelevant, if we assumed that  $RO$  were correct.

For example, the relevance value of the proposition “Vaccinations are dangerous” ( $AS_1$ ), which is an acceptability support for the main claim “Children should not be vaccinated,” is a measure of the extent to which “Vaccinations are dangerous” would rationally compel us to accept “Children should not be vaccinated,” if we assumed that the proposition “Vaccinations are dangerous” were correct. If it is correct, this would greatly enhance the acceptability of the main claim that children should not be vaccinated. This is why the acceptability support “Vaccinations are dangerous” receives a high relevance score.

For another example, let us consider the proposition “A study in the *BMJ* shows that children given MMR vaccinations were much more likely to become autistic” ( $AS_1AS_1AS_1$  in Row 4). If we assume the proposition to be correct, then we should be rationally compelled to also accept the proposition it is aimed at (“Vaccinations cause autism,”  $AS_1AS_1$ , in Row 3). In other words, when evaluating the relevance value of a proposition, we assign the value independently of whether or not the proposition is, in fact, true. The acceptability of the proposition is evaluated separately, and it is taken into consideration during the determination of the proposition’s rational force value.

Just as with acceptability values, the assignment of relevance values is influenced by higher-level relevance supports, if there are any. For example, let us consider the relevance support “*BMJ* is one of the world’s most prestigious journals” ( $RS_1AS_1AS_1AS_1$  in Row 5). This proposition was assigned a very high acceptability value. It also had high relevance with regard to the proposition “A study in the *BMJ* shows that children given MMR vaccinations were much more likely to become autistic” ( $AS_1AS_1AS_1$  in Row 4), resulting in its high rational force. The high rational force value of this relevance support in turn makes the relevance value of the proposition  $AS_1AS_1AS_1$  high.

### **Rational Force Values**

The *rational force* value of a proposition—that is the strength with which it influences the proposition it is aimed at—is fully determined by its acceptability and relevance values. According to Björnsson et al. (2009), “for a

proposition in an argument to be good—for it to have high rational force—it is required that it is both *acceptable* and *relevant*” (p. 12, as translated by the authors). Björnsson et al. (2009) provided general guidelines for combining acceptability and relevance values to come up with the rational force:

- A proposition has the highest possible rational force when both acceptability and relevance are the highest possible.
- If the relevance of a proposition is not the highest possible, the rational force decreases to an at least equivalent degree. If there is an entire lack of relevance, there is also an entire lack of rational force.
- The same applies to acceptability. If the acceptability of a proposition is not the highest possible, the rational force decreases to an at least equivalent degree.
- If both acceptability and relevance are lower than the highest possible, then the rational force “inherits” both these weaknesses. If the acceptability pulls down the rational force a bit, the relevance pulls it down further. (Björnsson et al., 2009, pp. 44–45, as translated by the authors)

For instance, although the relevance of the proposition “A study in the *BMJ* shows that children given MMR vaccinations were much more likely to become autistic” ( $AS_1AS_1AS_1$  in Row 4) is high, its acceptability is very low, resulting in a very low value for the rational force. We further developed the guidelines proposed by Björnsson et al. (2009) into a matrix presented in Figure 3. The matrix shows how to combine acceptability and relevance values assigned using a 5-point ordinal scale to determine a rational force of a given proposition. We relied on this matrix to determine the rational force values in Figure 2 and in Appendix A.

From Figure 3, we can see that if a proposition is low in *either* acceptability or relevance, its rational force is also low, making it weak in relation to its specific aim. However, we also need to consider that the aims are different for the four types of propositions: acceptability support, relevance support, acceptability objection, and relevance objection (see Figure 1). According to prior scholarship (Backman et al., 2012; Björnsson et al., 2009), the rational forces can be interpreted in relation to the different aims as follows:

- a. The *rational force* of an acceptability support *AS* for a proposition *P* is a measure of the extent to which *AS* would rationally compel us to accept *P*.
- b. The *rational force* of an acceptability objection *AO* against a proposition *P* is a measure of the extent to which *AS* would rationally compel us to reject *P*.

		Acceptability				
		Very high	High	Medium	Low	Very low
Relevance	Very high	Very high	High	Medium	Low	Very low
	High	High	Medium	Low	Low	Very low
	Medium	Medium	Low	Low	Very low	Very low
	Low	Low	Low	Very low	Very low	Very low
	Very low	Very low	Very low	Very low	Very low	Very low

**Figure 3.** Matrix of rational force values based on relevance and acceptability values, assigned with a 5-point ordinal scale.

- c. The *rational force* of a relevance support *RS* for a proposition *P* is a measure of the extent to which *AS* would rationally compel us to view *P* as relevant.
- d. The *rational force* of a relevance objection *RO* against a proposition *P* is a measure of the extent to which *AS* would rationally compel us to view *P* as irrelevant.

By weighing together the rational forces of all first-level propositions, we can assign a single *acceptability* value to the main claim. We use the term “acceptability,” rather than “rational force,” to describe the final judgment about the main claim. This is because the main claim is, by definition, a proposition that is not directed toward or against another proposition, and thus it cannot be attributed a relevance value. Since the rational force value is the product of the relevance and acceptability values, it does not apply to the main claim. Neither can there be any relevance supports or objections directed at the main claim, since the main claim cannot be relevant or irrelevant.

Note that the acceptability value of the main claim is determined by weighing together the rational forces of the first-level acceptability supports *and* objections. Gardelli et al. (2019) developed detailed and theoretically based procedures for taking into account the rational force values of these first-level propositions, but they are beyond the scope of this article. Instead, we present several basic rules, proposed by Björnsson et al. (2009, p. 49, as translated by the authors), that can guide the process of weighing together the rational forces of supports and objections:

- The total strength of the propositions *supporting* a certain proposition *increases* the more there are and the stronger they are.
- The total strength of the propositions *opposing* a certain proposition *increases* the more there are and the stronger they are.
- Propositions that are *supporting and opposing* a main claim and that *have equal strength cancel each other out*.

For example, in the anti-vaccination text, the only two first-level propositions are acceptability supports ( $AS_1$  and  $AS_2$ , see Rows 2 and 20 in Appendix A), and both of these have very low rational force values. Our RFM analysis shows that the accumulated acceptability of the first-level proposition “Vaccinations are dangerous” ( $AS_1$ ) is very low, and the same goes for the second first-level proposition, “There is now a strong expert opposition to vaccination” ( $AS_2$ ). Even though both of these propositions have high relevance, their rational forces are very low, because of their very low acceptability. Together, these first-level acceptability supports fail in providing strong reasons to believe that the main claim is correct.

To conclude, based on the RFM analysis of the entire text shown in Appendix A, the acceptability value for the main claim, “Children should not be vaccinated” (shown in row 1, under “A”) is very low. Our RFM analysis reveals that although the anti-vaccination argument has a complex structure, contains many desirable argument elements, and would score high using structural approaches, it fails to provide acceptable and relevant reasons to support the main claim.

### *Using Numeric Values*

In several publications on the RFM (e.g., Björnsson et al., 2009; Gardelli et al., 2019), researchers have proposed analytic schemes that use numbers between 0 and 1 to assign acceptability and relevance values. Working with numeric values adds clarity and precision to the RFM analysis, but it also increases the complexity and other challenges of the model and may present problems with the reliability of values assigned by different scorers.

Detailed discussion of the use of numeric values is beyond the scope of this article, but we include Figures 4 and 5 to demonstrate the potential of the RFM for quantitative analysis. In Figure 4, we provide two alternative methods of converting verbal values we used for acceptability and relevance in Figure 2 to numeric intervals. One of the methods relies on a linear transformation and another on nonlinear transformation. Although the linear transformation is more accessible and easier to apply, the nonlinear version is

Values	Linear numerical score intervals	Non-linear numerical score intervals
Very high	0.8 to 1	0.95 to 1
High	0.6 to < 0.8	0.8 to < 0.95
Medium	0.4 to < 0.6	0.2 to < 0.8
Low	0.2 to < 0.4	0.05 to < 0.2
Very low	0 to < 0.2	0 to < 0.05

**Figure 4.** Conversion of a verbal scale to numeric intervals.

		Acceptability				
		0.9	0.7	0.5	0.3	0.1
Relevance	0.9	0.81	0.63	0.45	0.27	0.09
	0.7	0.63	0.49	0.35	0.21	0.07
	0.5	0.45	0.35	0.25	0.15	0.05
	0.3	0.27	0.21	0.15	0.09	0.03
	0.1	0.09	0.07	0.05	0.03	0.01

**Figure 5.** Rational force values based on the RFM Multiplication model (Björnsson et al., 2009).

more theoretically appropriate and functions better at more advanced levels of analysis.

In Figure 5, we demonstrate a “Multiplication model” proposed by Björnsson et al. (2009) for combining relevance and acceptability values. According to this model, the rational force of a proposition in an argument is the product of the relevance and acceptability values. For example, after assigning relevance and acceptability values ranging from 0 to 1 for a proposition in an argument (for instance,  $R=0.6$  and  $A=0.8$ ), we multiply these to get the proposition’s rational force ( $0.6 \times 0.8=0.48$ ). In Figure 5, we give more examples based on the Multiplication model, where we use the mean values for each relevance and acceptability interval in the linear numerical score intervals column in Figure 4.

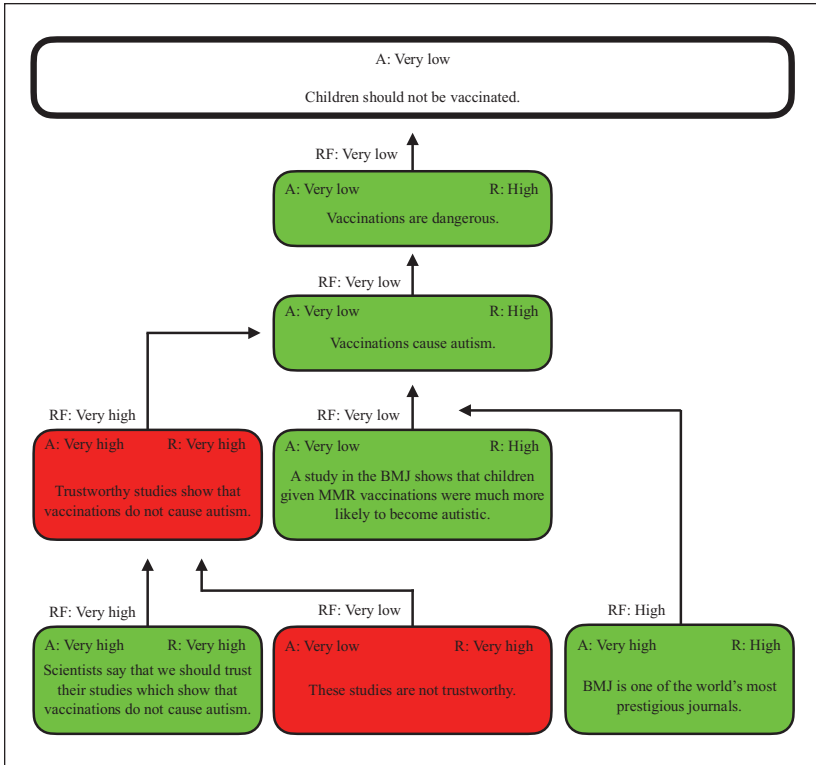


## Affordances and Challenges of RFM-Based Assessments

We argue that the RFM has the potential to improve the assessment of student arguments in several ways. First, the RFM supports evaluation of the content of each proposition, based on acceptability and relevance, two criteria that are widely considered to be fundamental markers of argument strength (Ennis, 2003; Govier, 2010; Nielsen, 2013). Second, the RFM provides a comprehensive, detailed, and transparent account of all propositions used in the argument and it shows the relations between those propositions. Third, the RFM yields a single value that captures the rational force of the argument as a whole with regard to its main claim, showing the extent to which the author succeeds in providing support for the main claim.

For researchers, the diagnostic information that is made available with RFM-based assessments could assist with designing interventions to develop students' competency in argumentation. As previously noted, our current understanding of argumentation development and the effectiveness of related instructional interventions is largely based on student outcome measures that rely on structural approaches (Berland & McNeill, 2010; Butler & Britt, 2011; Osborne et al., 2004; Reznitskaya et al., 2008; Yeh, 1998). This understanding could be enhanced, and possibly revised, if researchers assess student performance using the RFM, a framework that evaluates the acceptability and relevance of *every* proposition in an argument. By having an explicit focus on these two criteria, researchers may be better equipped to study conditions that help students acquire a more robust knowledge of argumentation and develop dispositions to value a truth-seeking process.

For teachers, using assessments based on the RFM could help enhance their understanding of argumentation and revise their instructional practices. In preparation for a lesson, a teacher could construct a graphic representation of an excerpt of an argument, such as that shown in Figure 6, and assign acceptability and relevance values to different propositions. In this way, the levels of various propositions readily become apparent (vertically), as do the differences between types of propositions—acceptability supports and objections are pointed directly toward other boxes of propositions, whereas relevance supports and objections are pointed toward links between different boxes. Whether a proposition serves as a support or an objection could be captured by the color of the box with, for example, supports shown in green and objections in red. Using the graphic representation, a teacher could discuss argument evaluation with students and engage them in reflection on why certain propositions have lower acceptability values and relevance values than others.



**Figure 6.** An example of visual representation of an excerpt from the anti-vaccination text.

Similarly, students could construct their own graphic representations of arguments in preparation for writing a paper or critiquing a peer's paper. This would further promote reflection on the epistemic criteria that underlie quality arguments. For students, the key understanding afforded by the RFM is that, because the rational force value is the product of the relevance and acceptability values, a proposition must be *both* acceptable and relevant in order for an argument to be strong.

Discussion of an RFM-based assessment of an argument could also help teachers and students reflect on the epistemic criteria that underlie knowledge justification in different disciplinary contexts. For example, the acceptability of a literary interpretation can be supported with details from the text; the acceptability of historical conclusions can be justified by reference to historical documents and artifacts; and that of scientific claims by systematic

experimentation, observation, and interpretation of data. Having discussions about acceptability and relevance of propositions in different disciplines could help teachers and their students consider the affordances and limitations of different epistemologies and promote the development of epistemic cognition aligned with disciplinary standards of knowledge construction (Barzilai & Chinn, 2020; Bråten et al., 2017; Goldman et al., 2016).

Although we are encouraged to envision new possibilities with RFM-based assessments, we also recognize the need for various modifications to make RFM applications feasible in a typical classroom. In our current work, we are exploring such modifications, including those shown in Figures 6 and 7. These figures provide examples of ways of translating the RFM for use in educational settings. A guiding principle is to preserve essential features of the RFM, such as distinguishing between the four types of propositions and different levels of propositions in arguments, as well as evaluating the acceptability and relevance of each proposition, while also making the RFM-based assessments accessible, useful, and feasible for teachers and their students.

For example, we plan to streamline the procedures for deriving acceptability, relevance, and rational force values by having teachers use a 3-point scale, with high, medium, and low values, as shown in Figure 7, instead of the 5-point scale described earlier in this article. In addition, we aim to clarify complex concepts using prompts such as “why believe” for acceptability and “why care” for relevance (see the column labeled “Type” in Figure 7 for an example). These modifications are in line with previous incorporations of everyday language in instructional RFM materials, such as “Why?” and “How do you know?” (acceptability questions) and “So what?” and “What does this have to do with anything?” (relevance questions) (Backman et al., 2012, p. 88, as translated by the authors).

To help teachers assess argumentative writing with a large number of students, we may need to design a scoring rubric with analytic categories representing key features of the RFM (e.g., acceptability, relevance). This rubric will be informed by the RFM, but it would not require teachers to engage in formal evaluation of each proposition in a student’s essay.

## Implications for Future Research

In this article, we presented a critical review of structural approaches to evaluating argument quality. We then described another framework, the RFM, and explained how RFM-based assessments can support a theoretically sound, systematic, and transparent analysis of naturally occurring arguments based on the criteria of acceptability and relevance. We suggest that explicit focus on these criteria, which is largely missing from current assessment

Level	Type	Student Wording	Meaning	Evaluation		
				Acceptability	Relevance	Rational force
	Main claim	"Children should not be vaccinated."		Low		
1	Why believe	"This means that vaccinations are dangerous ..."	Vaccinations are dangerous.	Low	High	Low
2	Why believe	"... and cause autism."	Vaccinations cause autism.	Low	High	Low
3	Why believe	"A study published by Wakefield and colleagues in British Medical Journal ... showed that children given MMR vaccinations were much more likely to become autistic."	A study in the BMJ shows that children given MMR vaccinations were much more likely to become autistic.	Very low	High	Very low
4	Why care	"... one of the world's most prestigious journals ..."	BMJ is one of the world's most prestigious journals.	Very high	High	High

**Figure 7.** Modification to support RFM-based assessments for teachers.

practices, would help engage students and their teachers in discussions of why some claims are more trustworthy than others, thus promoting deeper understanding of epistemic criteria and processes that underlie disciplinary knowledge construction. Such an understanding should, in turn, result in more robust learning of complex topics and lead to the development of epistemic cognition that reflects justification standards and practices used by professionals in academic disciplines (Bråten et al., 2017).

Despite the potential of the RFM for assessment and instruction, we also recognize the need for further theoretical developments and empirical research on the use of the RFM in educational contexts. For example, additional analytic categories could be applied to propositions in an RFM-based analysis. These categories might include a *clarity* value for each proposition, as well as other diagnostic indicators for the entire argument, such as *width*, or the number of distinct propositions of the same level; *depth*, or the number of levels of propositions; and *alternatives*, or the number of opposing propositions.

Let us first consider the category of *clarity*, as an example. When a text written by a student is unclear, it may cause problems both in the Descriptive and the Evaluative phases. For example, in the Descriptive phase, there may be interpretative difficulties in both the propositional reconstruction and the argument reconstruction phases. In a propositional reconstruction, ambiguous sentences may permit multiple interpretations and require substantial changes in the original wording to articulate the precise meaning of the proposition. Unclear texts also cause difficulties in the Evaluative phase. As Björnsson et al. (2009) explained, "When one determines the rational force in different arguments, one soon notices that it is important that both the main claim and the propositions are clearly formulated. Unclearly stated propositions have unclear rational force and it is difficult to evaluate the rational

force of propositions supporting unclearly stated main claims” (p. 51, as translated by the authors). In the context of classroom instruction, the need to make substantial changes to the original wording might indicate a weakness in a student’s argumentative writing. For example, two students might produce argumentative essays that, when reconstructed, turn out to be the same, but where larger modifications are needed for one of the essays. Adding an analytic dimension that represents *clarity* would help to create a more nuanced and comprehensive account of differences in students’ performance.

Researchers may also be interested in other features of student’s arguments. For example, an analytic category called “*width*” could help to reveal the number of distinct propositions at the same level, indicating how broad and comprehensive an argument is. This horizontal aspect of an argument is unnecessary in the judgment of argument quality if the author has shown the main claim to be correct through a single first-level acceptability support proposition with a very high rational force value. However, when an argument is about a contestable question, this is rarely the case, and there may be a need for several first-level acceptability supports that together make the main claim more acceptable. The requirement for *width* of an argument is therefore dependent on, among other things, the rational force scores of first-level acceptability supports.

On the other hand, some arguments may be characterized by several weak first-level acceptability supports but lack further elaboration of them. These arguments would be *wide*, but not *deep*. The *depth* of an argument could be another RFM-based category that reflects the number of levels of propositions, indicating how detailed and elaborated the argument is. As with argument *width*, argument *depth* is not a necessary requirement for a strong argument. The requirement for *depth* is contingent on the RFM scores of lower-level propositions, such as first-level acceptability supports. That is, a first-level proposition with a very high rational force value makes further elaboration unnecessary. The construct of argument *depth* needs to be investigated further to provide clear guidelines for valid assessment.

*Alternatives* is another useful category that could provide diagnostic information about the number of *first-level* acceptability objections, as well as the number of acceptability objections or relevance objections against first-level or higher-level supporting propositions. Consideration of alternatives is a desirable feature of students’ arguments (Asterhan & Schwarz, 2016; Kuhn & Moore, 2015; Osborne, 2010). It has been previously advocated by Næss (1959), who stated, “Under all conditions, we must expect that, when we have reached a conclusion—whether based on thorough or superficial considerations—we lose the ability to consider counter-arguments to some

extent. Therefore, we must spend more energy on counter-arguments than on further pro-arguments that support an already presented conclusion. Usually, we do the opposite. We collect new pro-arguments but overlook or forget new counter-arguments” (Næss, 1959, p. 81, as translated by the authors). Further research is needed to determine how to account for consideration of alternatives when using the RFM for assessment purposes. For instance, students’ *alternative* scores should increase if they incorporated relevant objections to the main claim and responded with acceptable and relevant objection, while the scores should not increase if students used irrelevant objections.

Another direction for further theoretical development in adapting the RFM to an educational context is to explore productive ways of combining the RFM with analytic strategies from other frameworks, such as Walton’s argument schemes and dialogue types (Walton, 1996a; Walton et al., 2008). During an RFM analysis, each proposition is assigned acceptability and relevance values, but different scorers may rely on various considerations when assigning these values. This could be considered as a strength of the RFM, as it provides placeholders to be instantiated by discipline-specific epistemic criteria and practices. At least some of these criteria and practices are described by the argument schemes developed by Walton and colleagues (Walton, 1996a; Walton et al., 2008). For example, Walton’s scheme “argument from expert opinion” and the accompanying set of critical questions (Walton et al., 2008) can be used to help with the assignment of relevance values in the anti-vaccination text. For instance, the critical question “Is *E* an expert in the field that [proposition] *A* is in?” can alert the scorer to the fact that personal knowledge of a parent generally lies outside the scope of medical expertise. Accordingly, the proposition “The parent feels that the children would be better off without vaccination” ( $AO_1AO_1AS_1$  in Appendix A) is given a very low relevance value, minimizing the proposition’s rational force. Thus, Walton’s argument schemes may inform judgments about relevance during an RFM analysis.

Importantly, the choice of normative argument schemes and related critical questions could be informed by another useful framework developed by Walton—his typology of dialogue types. According to Walton (1998), argumentation happens in different contexts, or dialogue types, which have “distinctive goals and methods used by the participants to achieve these goals” (p. 3). Walton identified six dialogue types—inquiry, persuasion, negotiation, information-seeking, deliberation, and quarrel—and described normative protocols for engagement within each type. For example, *inquiry dialogue* is a collaborative process with the goal of working toward the most reasonable conclusion. By contrast, *persuasion dialogue* is a competitive process aimed at defending chosen positions and undermining opponents’ arguments. This

difference in dialogue types is important because the relevance of a proposition could be determined by the degree to which it contributes to the dialogue's goal. Future research should examine how the concept of relevance is understood in relation to different dialogue types.

In addition to theoretical developments, the field would benefit from additional empirical studies that examine the psychometric properties of RFM-based assessments. For example, we need to establish whether different scorers can reliably engage in propositional and argument reconstruction and assign consistent acceptability and relevance values. We also need to use both correlational and qualitative methodologies to better understand how RFM-based scores compare to those derived from other analytic and holistic scoring frameworks, including assessments based on largely structural approaches. To further validate the functioning of the RFM-based scores, we need to examine how argumentation skills assessed with the RFM relate to other variables, such as disciplinary knowledge, epistemic cognition, and demographic characteristics.

We also need to study how teachers might use RFM-based assessments in the classroom. For example, it is important to identify the ways in which RFM-based assessments shape instruction in argumentative writing and inform individual feedback provided to their students. Also, considering that teachers' epistemic cognition affects their practices and student learning (Brownlee et al., 2011; Greene et al., 2016; Maggioni & Parkinson, 2008), we need to develop an evidence-based understanding about the instructional value of RFM-based assessments for teachers and examine how these assessments can support teacher learning about argumentation and epistemic criteria that underlie disciplinary judgments.

Finally, we acknowledge that the procedures presented in the paper may seem too complicated for assessment and instruction in educational contexts. However, our experience is assuredly in line with Björnsson's et al. (2009) conclusion that consistent use of the RFM "trains you to notice the relations between the propositions in a text, even when you are not using pencil and paper or computer and makes it easier to . . . quickly evaluate the [argument's] rational force" (p. 56, as translated by the authors)." Under the assumption that all arguments are reducible to only four proposition types, together with a main claim, the RFM offers a very powerful yet parsimonious framework for evaluating argument quality. Our hope is that engagement in RFM-based analysis, in whatever form it takes, will help teachers, students, and others to consider carefully the key qualities of propositions that make them more or less trustworthy and thereby "cultivate deep-seated and effective habits of discriminating tested beliefs from mere assertions, guesses, and opinions" (Dewey, 1910, p. 28).

## Appendix

Prop. No.	Argument reconstruction	Verbatim	Propositional reconstruction	Evaluation		
				A	R	RF
1	T	Children should not be vaccinated.	VL			
2	AS <sub>1</sub>	"This means that vaccinations are dangerous . . ."	VL	H	H	VL
3	AS <sub>1</sub> AS <sub>1</sub>	" . . . and cause autism."	VL	H	H	VL
4	AS <sub>1</sub> AS <sub>1</sub> AS <sub>1</sub>	"A study published by Wakefield and colleagues in <i>British Medical Journal</i> . . . showed that children given MMR vaccinations were much more likely to become autistic."	VL	H	H	VL
5	RS <sub>1</sub> AS <sub>1</sub> AS <sub>1</sub> AS <sub>1</sub>	" . . . one of the world's most prestigious journals . . ."	VH	H	H	H
6	RS <sub>1</sub> RS <sub>1</sub> AS <sub>1</sub> AS <sub>1</sub> AS <sub>1</sub>	"This means that . . ."	H	VH	VH	H
7	AO <sub>1</sub> AS <sub>1</sub> AS <sub>1</sub>	" . . . studies . . . show that vaccinations do not cause autism . . ."	VH	VH	VH	VH
8	AS <sub>1</sub> AO <sub>1</sub> AS <sub>1</sub> AS <sub>1</sub>	"Scientists say that we should trust their studies . . ."	VH	VH	VH	VH

(continued)



## Appendix (continued)

Prop. No.	Argument reconstruction	Verbatim	Propositional reconstruction	Evaluation		
				A	R	RF
9	AO <sub>1</sub> AO <sub>2</sub> AS <sub>1</sub>	"... and should not be taken seriously."	These studies are not trustworthy.	VL	VH	VL
10	AS <sub>1</sub> AO <sub>1</sub> AO <sub>2</sub> AS <sub>1</sub> AS <sub>1</sub>	"... they are quite biased ..."	The scientists behind the studies were quite biased.	VL	H	VL
11	AS <sub>1</sub> AS <sub>1</sub> AO <sub>1</sub> AO <sub>2</sub> AS <sub>1</sub> AS <sub>1</sub>	"... the scientists who conducted them expected the vaccinations to work..."	The scientists behind the studies expected the vaccinations to work.	H	VL	VL
12	RS <sub>1</sub> AS <sub>1</sub> AS <sub>1</sub> AO <sub>1</sub> AO <sub>2</sub> AS <sub>1</sub> AS <sub>1</sub>	"... so ..."	If scientists who conduct vaccination studies expect the vaccinations to work, then they are quite biased.	VL	VH	VL
13	AO <sub>1</sub> AS <sub>1</sub>	"... my children will be safer if given a vaccination ..."	The parent's children will be safer if given a vaccination.	H	M	L
14	AS <sub>1</sub> AO <sub>1</sub> AS <sub>1</sub>	"Doctors say that ..."	Doctors hold that the parent's children will be safer if given a vaccination.	VH	H	H
15	RO <sub>1</sub> AS <sub>1</sub> AO <sub>1</sub> AS <sub>1</sub>	"... they do not have personal knowledge of my children."	The doctors do not have personal knowledge of the parent's children.	VH	VL	VL
16	RS <sub>1</sub> RO <sub>1</sub> AS <sub>1</sub> AO <sub>1</sub> AS <sub>1</sub>	N/A	Personal knowledge is important in determining whether or not vaccination is healthy for another person or not.	VL	VH	VL
17	AO <sub>1</sub> AO <sub>1</sub> AS <sub>1</sub>	"... I know in my heart that they are better off without taking that terrible risk."	The parent feels that the children would be better off without vaccination.	VH	VL	VL

(continued)

## Appendix (continued)

Prop. No.	Argument reconstruction	Verbatim	Propositional reconstruction	Evaluation		
				A	R	RF
18	$RS_1AO_1AO_1AS_1$	"I do . . ."	The parent has personal knowledge of the children.	VH	VL	VL
19	$RS_1RS_1AO_1AO_1AS_1$	N/A	Personal knowledge is important in determining whether or not vaccination is healthy for another person or not.	VL	VH	VL
20	$AS_2$	" . . . there is now a strong expert opposition to vaccination."	There is now (at the time when the text was written) a strong expert opposition to vaccination.	VL	VH	VL
21	$AS_1AS_2$	" . . . there are more and more experts who are coming around and opposing vaccinations and debating the scientific establishment."	There are more and more experts who are coming around and opposing vaccinations and debating the scientific establishment.	VL	M	VL
22	$AS_1AS_1AS_2$	"Now lots of US television programs are shedding light on this issue . . ."	Now (at the time when the text was written) lots of US television programs are shedding light on this issue.	M	L	VL
23	$RS_1AS_1AS_1AS_2$	" . . . which shows that . . ."	If lots of US television programs are shedding light on the vaccination issue, then there are more and more experts who are coming around and opposing vaccinations and debating the scientific establishment.	L	VH	L

## Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

## References

- Asterhan, C. S. C., & Schwarz, B. B. (2016). Argumentation for learning: Well-trodden paths and unexplored territories. *Educational Psychologist, 51*(2), 164-187. <https://doi.org/10.1080/00461520.2016.1155458>
- Backman, Y., Gardelli, T., Gardelli, V., & Persson, A. (2012). *Scientific thinking tools: A base for academic studies*. Studentlitteratur.
- Barzilai, S., & Chinn, C. A. (2020). A review of educational responses to the “post-truth” condition: Four lenses on “post-truth” problems. *Educational Psychologist, 55*(3), 107-119. <https://doi.org/10.1080/00461520.2020.1786388>
- Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education, 94*(5), 765-793. <https://doi.org/10.1002/sce.20402>
- Björnsson, G., Kihlbom, U., & Ullholm, A. (2009). *Argumentation analysis: Skills for critical thinking*. Natur & kultur.
- Björnsson, G., Kihlbom, U., Tersman, F., & Ullholm, A. (1994). *Argumentationsanalys*. Natur och kultur.
- Bolsen, T., & Palm, R. (2022). Politicization and COVID-19 vaccine resistance in the U.S. *Progress in Molecular Biology and Translational Science, 188*(1), 81-100. <https://doi.org/10.1016/bs.pmbts.2021.10.002>
- Bråten, I., Muis, K., & Reznitskaya, A. (2017). Teachers’ epistemic cognition in the context of dialogic practice: A question of calibration? *Educational Psychologist, 52*(4), 253-269. <https://doi.org/10.1080/00461520.2017.1341319>
- Brownlee, J., Schraw, G., & Berthelsen, D. (Eds.). (2011). *Personal epistemology and teacher education*. Routledge.
- Butler, J. A., & Britt, M. A. (2011). Investigating instruction for improving revision of argumentative essays. *Written Communication, 28*(1), 70-96.
- Chinn, C. A., Duncan, R. G., Hung, L. C.-C., & Rinehart, R. W. (2016, April). *Epistemic criteria and reliable processes as indicators of argument quality in science students’ argumentation* [Conference paper]. Annual Meeting of the American Educational Research Association, Washington, DC.
- Crosswhite, J. (2018). Toward truth. *Philosophy & Rhetoric, 51*(4), 368-391.
- De La Paz, S., Ferretti, R., Wissinger, D., Yee, L., & MacArthur, C. (2012). Adolescents’ disciplinary use of evidence, argumentative strategies, and organizational structure in writing about historical controversies. *Written Communication, 29*(4), 412-454. <https://doi.org/10.1177/0741088312461591>

- Dewey, J. (1910). *How we think*. D. C. Heath & Co., Publishers.
- Dong, T., Anderson, R. C., Kim, I., & Li, Y. (2008). Collaborative reasoning in China and Korea. *Reading Research Quarterly*, 43(4), 400-424. <https://doi.org/10.1598/RRQ.43.4.5>
- Ennis, R. (2003). Critical thinking assessment. In D. Fasko (Ed.), *Critical thinking and reasoning* (pp. 293-313). Hampton Press, Inc.
- Erduran, S. (2018). Toulmin's argument pattern as a "horizon of possibilities" in the study of argumentation in science education. *Cultural Studies of Science Education*, 13(4), 1091-1099. <https://doi.org/10.1007/s11422-017-9847-8>
- Erduran, S., Simon, S., & Osborne, J. (2004). TAPPING into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, 88(6), 915-933. <https://doi.org/10.1002/sce.20012>
- Ferretti, R. P., Lewis, W. E., & Andrews-Weckerly, S. (2009). Do goals affect the structure of students' argumentative writing strategies? *Journal of Educational Psychology*, 101(3), 577-589.
- Føllesdal, D., Walløe, L., & Elster, J. (1990). *Argumentasjonsteori, språk og vitenskapsfilosofi*. Universitetsforlaget AS.
- Gardelli, V., Backman, Y., & Gardelli, T. (2019). *Argumentation analysis: New theory and new applications* [Paper presentation]. Filusofia, Luleå University of Technology.
- Goldman, S. R., Britt, M. A., Brown, W., Cribb, G., George, M., Greenleaf, C., Lee, C. D., Shanahan, C., & Project, R. (2016). Disciplinary literacies and learning to read for understanding: A conceptual framework for disciplinary literacy. *Educational Psychologist*, 51(2), 219-246. <https://doi.org/10.1080/00461520.2016.1168741>
- Govier, T. (2010). *A practical study of argument* (7th ed.). Wadsworth Publishing Company.
- Greene, J. A., Sandoval, W. A., & Bråten, I. (Eds.). (2016). *Handbook of epistemic cognition*. Routledge.
- Kavanagh, J., & Rich, M. D. (2018). *Truth decay: An initial exploration of the diminishing role of facts and analysis in American public life*. Rand Corporation.
- Kim, M., & Roth, W.-M. (2018). Dialogical argumentation in elementary science classrooms. *Cultural Studies of Science Education*, 13(4), 1061-1085. <https://doi.org/10.1007/s11422-017-9846-9>
- Kuhn, D. (2010). Teaching and learning science as argument. *Science Education*, 94(5), 810-824. <https://doi.org/10.1002/sce.20395>
- Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science*, 22, 545-552. <https://doi.org/10.1177/0956797611402512>
- Kuhn, D., & Moore, W. (2015). Argumentation as core curriculum. *Learning: Research and Practice*, 1(1), 66-78. <https://doi.org/10.1080/23735082.2015.994254>

- Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child development*, 74(5), 1245-1260. <https://doi.org/10.1111/1467-8624.00605>
- Maggioni, L., & Parkinson, M. (2008). The role of teacher epistemic cognition, epistemic beliefs, and calibration in instruction. *Educational Psychology Review*, 20(4), 445-461. <https://doi.org/10.1007/s10648-008-9081-8>
- McNeill, K. L. (2011). Elementary students' views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. *Journal of Research in Science Teaching*, 48(7), 793-823. <https://doi.org/10.1002/tea.20430>
- McNeill, K. L., & Knight, A. M. (2013). Teachers' pedagogical content knowledge of scientific argumentation: The impact of professional development on K-12 teachers. *Science Education*, 97(6), 936-972. <https://doi.org/10.1002/sc.21081>
- McNeill, K. L., & Krajcik, J. (2012). *Supporting grade 5-8 students in constructing explanations in science: The claim, evidence, and reasoning framework for talk and writing*. Pearson.
- Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. *Cognition and Instruction*, 14(2), 139-178.
- Næss, A. (1959). *Communication and argument: Elements of applied semantics*. Allen & Unwin.
- National Governors Association. (2010). *Common core state standards: Appendix A. research supporting key elements of the standards*. National Governors Association Center for Best Practices, Council of Chief State School Officers.
- Newell, G. E., Beach, R., Smith, J., VanDerHeide, J., Kuhn, D., & Andriessen, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly*, 46(3), 273-304. <https://doi.org/10.1598/RRQ.46.3.4>
- Nielsen, J. A. (2013). Dialectical features of students' argumentation: A critical review of argumentation studies in science education. *Research in Science Education*, 43(1), 371-393.
- Nussbaum, E. M., Dove, I. J., Slife, N., Kardash, C. M., Refika, T., & Vallett, D. (2019). Using critical questions to evaluate written and oral arguments in an undergraduate general education seminar: A quasi-experimental study. *Reading and Writing*, 32(6), 1531-1552. <https://doi.org/10.1007/s11145-018-9848-3>
- Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science*, 328(5977), 463-466. <https://doi.org/10.1126/science.1183944>
- Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, 41(10), 994-1020. <https://doi.org/10.1002/tea.20035>
- Oyler, J. (2019). Exploring teacher contributions to student argumentation quality. *Studia Paedagogica*, 24(4), 173-198. <https://doi.org/10.5817/SP2019-4-8>
- Partnership for 21st Century Skills. (2012). *A Framework for 21st Century Learning*. <http://www.p21.org/index.php>
- Postman, N. (1995). *The end of education: Redefining the value of school*. Knopf.

- Puri, N., Coomes, E. A., Haghbayan, H., & Gunaratne, K. (2020). Social media and vaccine hesitancy: New updates for the era of COVID-19 and globalized infectious diseases. *Human Vaccines & Immunotherapeutics*, 16(11), 2586-2593. <https://doi.org/10.1080/21645515.2020.1780846>
- Rapanta, C. (2022). Douglas Walton's contributions in education: A synthesis of theoretical and empirical research. *Informal Logic*, 42(1), 139-170.
- Rapanta, C., Garcia-Mila, M., & Gilabert, S. (2013). What is meant by argumentative competence? An integrative review of methods of analysis and assessment in education. *Review of Educational Research*, 83(4), 483-520. <https://doi.org/10.3102/0034654313487606>
- Reznitskaya, A., Anderson, R. C., Dong, T., Li, Y., Kim, I., & Kim, S. (2008). Learning to think well: Application of argument schema theory. In C. C. Block & S. Parris (Eds.), *Comprehension instruction: Research-based best practices* (pp. 196-213). Guilford Press.
- Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes*, 32(2-3), 155-175. [https://doi.org/10.1207/S15326950DP3202&3\\_04](https://doi.org/10.1207/S15326950DP3202&3_04)
- Reznitskaya, A., Kuo, L., Clark, A., Miller, B., Jadallah, M., Anderson, R. C., & Nguyen-Jahiel, K. (2009). Collaborative reasoning: A dialogic approach to group discussions. *Cambridge Journal of Education*, 39(1), 29-48. <https://doi.org/10.1080/03057640802701952>
- Reznitskaya, A., & Wilkinson, I. A. G. (2020). Measuring production and comprehension of written arguments in upper-elementary grades. *Studia Paedagogica*, 24(4), 63-84.
- Sampson, V., & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education*, 92(3), 447-472.
- Swedish National Agency for Education. (2021). *Curriculum for the compulsory school, preschool class and school-age educare*. Author.
- Toulmin, S. E. (1958). *The uses of argument*. Cambridge University Press.
- Walton, D. (1996a). *Argument structure: A pragmatic theory*. University of Toronto Press.
- Walton, D. (1996b). *Argumentation schemes for presumptive reasoning*. Erlbaum.
- Walton, D. (1998). *The new dialectic: Conversational contexts of argument*. University of Toronto Press.
- Walton, D. (2014). On a razor's edge: Evaluating arguments from expert opinion. *Argument and Computation*, 5(2-3), 139-159. <https://scholar.uwindsor.ca/crrarpub/25>
- Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation schemes*. Cambridge University Press.
- Yeh, S. (1998). Empowering education: Teaching argumentative writing to cultural minority middle-school students. *Research in the Teaching of English*, 33, 49-81.
- Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39(1), 35-62. <https://doi.org/10.1002/tea.10008>

**Author Biographies**

**Dr. Ylva Backman**, her research interests include argumentation theory, philosophy with children, philosophical dialogues in special needs education, well-being and dialogic education. She has many years experience in teaching and researching the RFM model.

**Prof. Alina Reznitskaya** has expertise in educational psychology, quantitative research methodology, and educational measurement. Her research interests include investigating the role social interaction plays in the development of argument literacy, designing measures of argumentation, and examining professional development efforts that help teachers improve the quality of argumentation during class discussions.

**Dr. Viktor Gardelli** has a background in Philosophy and Educational science, his research interests include ethics, the philosophy of education, philosophy with children, logic and argumentation theory, and technology in education. He has a particular interest for philosophical investigations concerning concepts such as truth, validity, relevance and rationality.

**Prof. Ian A. G. Wilkinson** has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers.