ARGUMENTATION AND SCIENTIFIC REASONING - AN EXPLORATION OF THEIR INTERRELATIONSHIP

Argumentation can be regarded as an important aspect of the nature of science which should be taught in school science. Current research indicates that students can be trained to argue, in particular, if specific prompts or scaffolds help students to identify when an argumentation is an expected pattern of reasoning. However, several situations require students to distinguish between evidence and conclusion or to evaluate claims and corresponding evidences without the opportunity to draw on prompts or scaffolds. Our research project aims to explore the interrelationship between argumentation and science learning. This broader research focus includes analyses of students' argumentations as they "naturally" occur during instruction that has a focus on activities which are assumed to promote conceptual understanding. In order to investigate students' argumentations in such situations, a distinction between argumentation and other types of reasoning is necessary. Video data from students of different age working in groups on physics and biology instruction are used to develop a coding manual which is based on Toulmin's Argument Pattern but distinguishes explicitly between argumentation and processes of clarifications and explanations. Results indicate that both argumentation and clarification occur in about up to 30% of the time spend with the instruction.

Claudia v. Aufschnaiter, Justus Liebig University Giessen, Institute of Physics Education
Jan Fleischhauer, Justus Liebig University Giessen, Institute of Physics Education
Christian Rogge, Justus Liebig University Giessen, Institute of Physics Education
Tanja Riemeier, Leibniz University Hannover, Biology Education

Background, Aims and Framework

In the last decade, an increasing body of research has addressed argumentation in science education. The learning of science and the learning about science are considered equally important to relevant science education (e.g., Millar & Osborne, 1998). As a consequence, several approaches have emerged aiming to promote and investigate students' argumentation in science classrooms (e.g., Chinn, O'Donnell, & Jinks, 2000; Jiménez Aleixandre & Pereiro Muñoz, 2005; Osborne, Erduran, & Simon, 2004; Schwarz, Neuman, Gil, & Merav, 2003). Common to most of these studies is the attempt to concentrate on specific prompts or scaffolds that enable and support students to argue about a scientific or socio-scientific issue. For instance, students should agree or disagree to a specific course of action (funding a zoo, allowing experiments with animals) or they should argue whether or not specific (theory-based) statements explain a particular phenomenon. Even though these approaches have shown to improve students' ability to argue in response to a given prompt (e.g., Osborne et al., 2004; Zohar & Nemet, 2002), it is yet unclear under which conditions will students engage in argumentation without prompts or scaffolds. However, several situations (in and out of school) require students to distinguish between evidence and conclusion or to evaluate claims and corresponding evidences without the opportunity to draw on prompts or scaffolds. Thus, it would be necessary to understand when and how students engage in argumentation even though they are not requested to argue.

Research reported in this paper aims to explore the interrelationship between students' argumentation and their conceptual development. This broader research focus includes analyses of students' argumentations as they "naturally" occur during instruction that has a focus on activities which are assumed to promote conceptual understanding. In order to investigate students' argumentation in such situations, a distinction between argumentation and other types of reasoning is necessary. Otherwise, processes in which students seek for a better understanding of a given phenomenon/situation are confused with processes in which students try to persuade someone of their belief. Whereas the former is related to the development of conceptual knowledge, the latter draws on some knowledge about evidences supporting a specific claim (e.g., von Aufschnaiter, Osborne, Erduran, & Simon, 2008). Though the need and suggestions for such distinction is present in some literature on argumentation (e.g., Mayes, 2000; Walton, 2006; Wright, 2002), research on students' argumentation often does not take explicitly into account which reasoning patterns count as an argumentation and which do not. Criteria such as the Toulmin Argument Patterns (e.g., Toulmin, 2003) can be applied to argumentation and other types of reasoning, for instance, explanations and processes of clarification. The review of the literature on argumentation indicates that a definition of what counts as a claim is needed in order to identify argumentation in settings which do not explicitly require students to argue (Houtlosser, 1998).

The framework outlined briefly indicates that one emphasis of the project is to develop criteria, from both a theoretical and empirical perspective, which allow to identify argumentation in "normal" settings. These criteria have been used to analyse in detail which conditions (individual, social and instructional) promote students' engagement in argumentation.

Design, Samples and Methods

Instruction used in the project is organized around argumentative activities taken from material developed at King's College London (e.g., Osborne et al., 2004). Four topics were chosen: electric circuits, heat transfer, light & shadow, and blood pressure. The setting consisted of three phases: a) instruction aiming to establish an understanding of the phenomena and concepts concerned with the task to argue about, b) the argument-task itself – taken from King's material, and c) a post-phase addressing contents deployed in the argumentative activity in order to investigate progress in students' understanding as a result of the argumentation in b). For each topic, these phases were distributed over two sessions, each lasting about 80 minutes. The first session and about two-thirds of the second session dealt with the instruction whereas the argument-task and the post-phase make up about 30 minutes. The design of the instruction was informed by the Model of Educational Reconstruction (e.g., Duit, Gropengießer, & Kattmann, 2005) and by results on students' learning processes in physics (e.g., von Aufschnaiter, 2006; von Aufschnaiter & von Aufschnaiter, 2003). All units developed were tested by a sample of 28 students from university who were working in groups of two to three students on one of the units. These students were studying social sciences and did not have a strong background in the topics covered by the units. Videoing of the processes informed us about necessary changes (such as for instruction that was not understood or experiments that did not work as expected). The main study consisted of two samples: 18 students from grade 8 (about 13 years old) 12 students from grade 11 (about 16 years old). The students worked in groups of three on two topics (Table 1).

Table 1
Distribution of Topics over Groups

	Year 8						*Year 11*			
Topic	8.1M	8.1F	8.2M	8.2F	8.3M	8.3F	11.1M	11.2F	11.3M	11.4F
electric circuits		x	x			x	x	x		
heat transfer	x			x	x				x	x
blood pressure		x		x	x				x	x
light and shadow	x		x			x	x	x		

All student activities were recorded on video. Investigating students' argumentation requires two steps. First, we had to identify sequences in which students were arguing (as distinct from those in which they were "just" discussing). Second, we had to investigate patterns of argumentation in these sequences. Even though research on argumentation suggests different (coding) procedures, several projects refer to Toulmin's Argument Pattern for analysis (Toulmin, 2003). Argumentations are initialized with claims being supported with justifications. These can either be data or warrants, which may be supported with backings. Furthermore, counter-claims and rebuttals are essential components of refuting one's argumentation while qualifiers restrict the validity of the claim (for more details see Erduran, Simon, & Osborne, 2004; v. Aufschnaiter et al., 2008). We, also, have been using Toulmin's Argument Pattern in order to analyse students' argumentations. However, although these patterns are widely used, we realized that detailed coding schemes were mainly absent from research reports. Especially, we found almost no information on how to identify exactly a "claim", especially in those cases in which the material does not explicitly ask for a specific claim. Furthermore, information on what (exactly) distinguishes a warrant from data or at least how to identify warrants seems to be rare and analyses differ between different groups. Hence, one main aim of our research has been to develop a coding schema and manual for both, the (video-based) identification and the (transcript-based) analysis of students' argumentations drawing on Toulmin's Argument Pattern.

Results

Coding Schema

In our research process, identifying students' argumentations and coding these argumentations were interrelated. We started the development of a coding schema by addressing those bits of the video data in which students argued about the argument-task. Here, identifying the claims was relatively easy as these were present with the material offered. For developing the schema, these data were transcribed (utterances and student activities) and then coded extensively, discussing how and why we applied the codes. From this process, we developed a coding schema that we can apply to transcribed data successfully (see Appendix A for the current version of the schema which is, in German, imbedded into a coding manual of about 15 pages length). Though we are still refining this schema, our inter-coder agreement already usually exceeds 75% between different raters (due to small sample sizes we have not yet calculated a Cohen's kappa).

One main criterion for identifying claims is an explicit focus on whether or not the claim includes an intention to persuade someone that the statement chosen, the prediction, or the activity proposed is valid/necessary (e.g., Mayes, 2000; Walton, 2006). Transcript 1, in contrast, illustrates that some dialogue could be coded as argumentation even though an intention to persuade is missing. Rather, the students in this example seem to seek for possible reasons for their observations. These reasons might appear to be claims (or counter-claims) but do not convey a specific belief. Thus, we would conclude, this dialogue is a clarification rather than an argumentation.

The students set-up an electric circuit with one battery (4.5 V) and two LEDs

	Transcript	**Possible coding for argumentation [annotation]**
S2:	Their [the LEDs] glowing is faint.	data [observation]
S1:	I cannot see them glow.	data [observation]
S2:	Swap the wires.	claim [swapping the wires should make a difference]
S1:	The batteries might be weak.	c-claim [the batteries are the problem]
	Why does only one LED work?	data [observation]
S3:	Is there a loose connection?	c-claim [the connection is the problem]
S1:	I got it. One LED is in the wrong direction.	c-claim [the position of the LED is the problem]

Transcript 1. Students discuss their observations and ideas for an electric circuit

Not as a surprise, we noted that coding transcripts is a time consuming activity which we could not apply to all our video data as this would require transcripts of about 64 hours of video. Therefore, we were aiming to find a coding procedure that could be applied to the video data directly (without transcripts) in order to identify at least those segments in which students argued (apart from all other activities). Using Toulmin's Argument Pattern for the videos (without transcripts) was not possible because they would require a detailed investigation of each single utterance. Therefore, we condensed the codes to a more general coding system including (1) claim & counter-claim, (2) claim & justification – complementary, (3) claim (& counter-claim) & justification – contradictory, (4) clarification, (5) other (either 1-3 or 4 which could not be identified clearly). With this coding schema at hand, we started to investigate all video data in order to identify and code especially those parts of the videos that were not related to an argument-task but, rather, comprised the "normal" learning contexts (which make up the majority of our data). Codes were applied at 10-secondintervals (time-based-sampling) using the software Videograph (Rimmele, 2007). Whenever we will present results on these coding procedures one have to bear in mind that all values which are based on these intervals are maximum values as typically the coded event is shorter than the interval.

So far, all videos are completely coded by at least one rater, about one-third of the videos were coded independently by a second rater. Again, intercoder agreement usually exceeds 80% with a Cohen's kappa of about 0.6 to 0.7. For both transcript-based coding and coding of the videos all disagreement between two raters as well as uncertainties of a single rater are resolved in detailed discussions amongst our research group. However, it should be noted that in about 5% of the instances we remain with two different ways to code the sequence depending on the interpretation of the sequence. At the video, such sequences would be condensed into category (5).

Coding the videos serves as a basis for both, quantitative analyses of applied codes and qualitative, transcript-based analysis of all sections that were coded with 1-4.

Argumentation and Clarification

We found that typically up to 30% of the total time was spent with argumentation or clarification (Figure 1a). Students from grade 11 spent about 10% more time for these types of reasoning but the ratio between argumentation and clarification is almost identical to the ratio from students of grade 8 (Figure 1b).

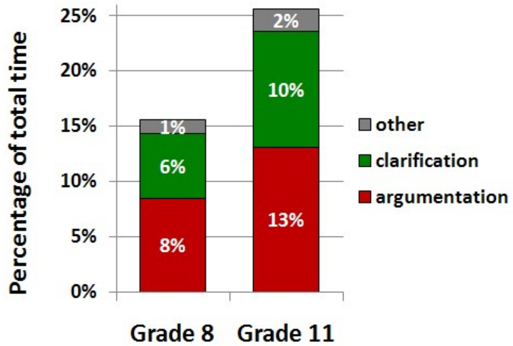

Figure 1a. Percentage of argumentation, clarification, and other of the total time spend with the units (based on 10-secondintervals)

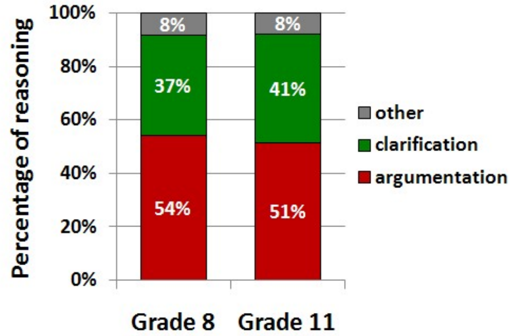

Figure 1b. Percentage of argumentation, clarification and other of all reasoning coded with 1-5 (based on 10-secondintervals)

The higher percentage of grade 11 students' argumentations and clarifications in Figure 1a is caused by both, more reasonings during the session but also these lasting longer. This is illustrated with Figure 2 which shows for grade 8 (dots) and grade 11 (squares) in how many intervals the students started a coherent argumentation or clarification (percentage of intervals shown on the y-axis) and the mean value for a coherent reasoning for the different groups (x-axis).

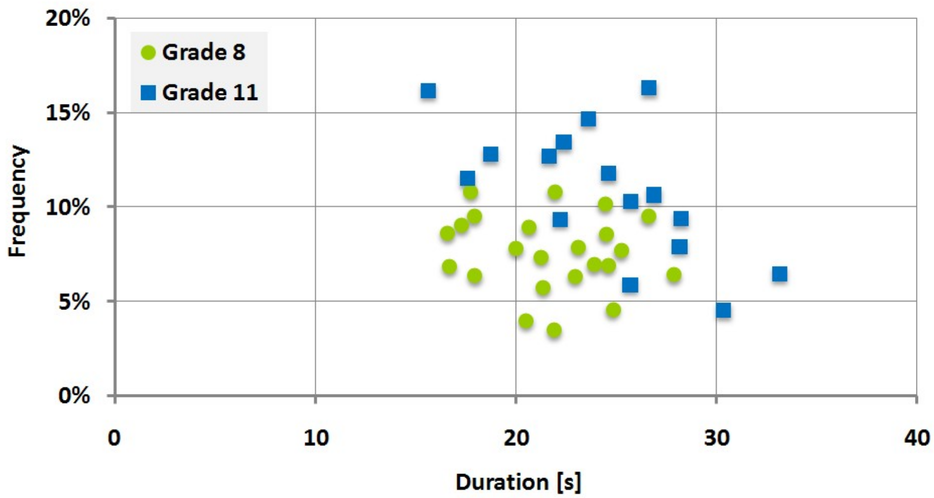

Figure 2. Mean duration of and frequency with which a coherent argumentation or clarification is started in all units for grade 8 and grade 11 students (results based on 10-secondintervals)

Further quantitative analysis revealed that longer reasonings developed by grade 11 students seem to be partly caused by these students developing more controversial arguments (code (3)) which last longer than "simple" claim-counterclaim (code (1)) or complementary arguments (code (2)). Even though the 11th graders developed extended arguments and clarifications, Figure 2 also illustrates that coherent reasonings did normally not exceed 30 seconds.

Although we have performed several quantitative analyses, looking for patterns in age, gender, or learning unit, we have not yet a clear picture of factors that might have caused differences. The variances inbetween groups of one cohort, inbetween cohorts, and between learning units (even between the two sessions of one unit) are high for both argumentation and clarification (Figures 2 and 3). However, our results clearly show that both argumentation and clarification occur in all three phases in our units, that is, students argue not only during the argument-tasks but also in the phases preceding and following these tasks.

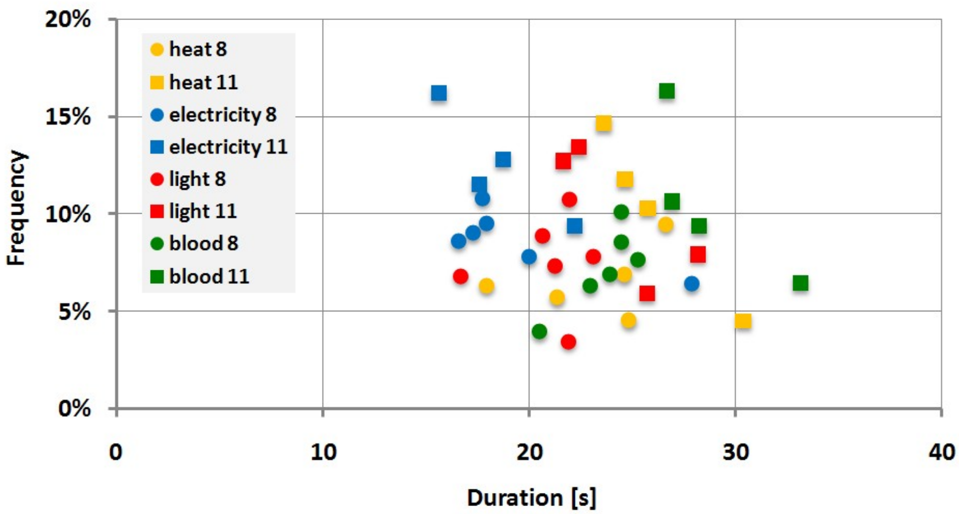

Figure 3. Mean duration of and frequency with which a coherent argumentation or clarification is started depending on unit, differentiated between first and second session, and grade (all results based on 10-secondintervals)

Our interpretation of the results gained so far is that the details of the processes matter. For instance, whether or not students understand the argument-task, can activate appropriate knowledge about evidences, have a dominant student in their group (maybe knowing "more" than the others), are successful with the experiments and/or find puzzling phenomena can have an impact on what kind of reasoning occurs, if any. Thus, understanding the interrelationship between argumentation and clarification (respectively the learning of science concepts) requires a qualitative in-depth analysis.

Preliminary results of the qualitative analyses indicate that students of grade 11 more often than those from grade 8 clarify evidences *before* they develop an argumentation. For instance, they elaborate and discuss the physics/biology background and meaning of evidences. Further analyses will investigate the parameters which have an impact on whether or not an argumentation is developed or a clarification or none. Especially, we expect that content specific knowledge, which we can classify in terms of its conceptual basis (see v. Aufschnaiter, 2006) plays an important role for the type of reasoning developed.

Conclusions and Implications

Theoretical distinctions considered, criteria used for the analyses, and empirical results are supposed to contribute to a precise definition of Toulmin's Argument Patterns in the context of science education. Furthermore, reasoning patterns which are strongly related to learning, such as processes of clarification and explanation, are interrelated to argumentation as a scientific activity. Thus, the project will contribute to a better understanding of how to promote students' argumentation as – maybe – asking for a clarification of evidences might be a better starting point than asking for a claim.

Acknowledgement: Research reported in this paper has been funded by the German Research Association (DFG AU 155/5-1,2). We are very grateful to Jonathan Osborne, Sibel Erduran, and Shirley Simon for sharing their material and thoughts on argumentation with us.

References

Chinn, C. A., O'Donnell, A. M., & Jinks, T. S. (2000). The structure of discourse in collaborative learning. *The Journal of Experimental Education, 69*(1), 77-97.

Duit, R., Gropengießer, H., & Kattmann, U. (2005). Towards science education research that is relevant for improving practice: The model of educational reconstruction. In H. E. Fischer (Ed.), *Developing standards in research on science education* (pp. 1-9). Leiden: Taylor & Francis.

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.

Houtlosser, P. (1998). Points of view. *Argumentation, 12*(3), 387-405.

Jiménez Aleixandre, M. P., & Pereiro Muñoz, C. (2005). Argument construction and change while working on a real environment problem. In K. T. Boersma, M. Goedhart, O. de Jong & H. Eijkelhof (Eds.), *Research and the quality of science education*.Dordrecht: Springer.

Mayes, G. R. (2000). Resisting explanation. *Argumentation, 14*(4), 361-380.

Millar, R., & Osborne, J. (1998). *Beyond 2000. Science education for the future*. London: Nuffield Foundation.

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.

Rimmele, R. (2007). Videograph. A multimedia-player for coding videos. Kiel: IPN.

Rogge, C., Aufschnaiter, C. v., Riemeier, T. & Fleischhauer, J. (2007). Kodierleitfaden für die Analyse von Argumentationsprozessen [Coding manual for analysing processes of argumentation]. Justus-Liebig-Universität und Leibniz Universität Hannover, 28.09.2007, unpublished.

Schwarz, B. B., Neuman, Y., Gil, J., & Merav, I. (2003). Construction of collective and individual knowledge in argumentative activity. *The Journal of the Learning Sciences, 12*(2), 219-256.

Toulmin, S. (2003). *The uses of argument* (updated edition). Cambridge: Cambridge University Press.

von Aufschnaiter, C. (2006). *Exploring the processes of students' development of physics concepts.* Paper presented at the Annual meeting of the National Association for Research in Science Teaching, San Francisco.

von Aufschnaiter, C., & von Aufschnaiter, S. (2003). Theoretical framework and empirical evidence on students' cognitive processes in three dimensions of content, complexity, and time. *Journal of Research in Science Teaching, 40*(7), 616-648.

von Aufschnaiter, C., Osborne, J., Erduran, S., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching, 45*(1), 101-131.

Walton, D. (2006). *Fundamentals of Critical Argumentation*. New York: Cambridge University Press.

Wright, L. (2002). Reasoning and explaining. *Argumentation, 16*(1), 33-46.

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.

Appendix A: Toulmin's Argument Pattern: Coding Schema for Transcripts (based on Rogge et al., 2007)

Category	*Description*	*Brief overview of indicators and hints*
claim (c)	Position-taking of a person (such as expressing a decision, formulation of an assertion or a prediction). From student's voice and intonation or the context it should be clear, that the student is convinced about the validity of his/her statement and intends to persuade someone. Typically, claims serve as the starting point of an argumentation.	• Explicit agreement to one or more given statements (e.g., "It has to be answer B!"). • A given statement is chosen which is supposed to refer to a particular phenomenon/event (e.g., "I am sure that statement B explains ..."). • A self-evident prediction is given for a phenomenon/event (e.g., predicting an effect that will occur in an experiment). • An activity is proposed to which no alternative approaches seem to be accepted (e.g., "No! Don't touch this."). ➢ Claims can be identified more easily in situations which ask for a position (e.g., within tasks that offer different statements one of which has to be chosen). Intonation and loudness of voice are helpful to identify claims apart of these contexts.
counter-claim (cc)	A counter-position or an alternative position is expressed to a claim.	• Disagreement with a claim, a written down statement, or a pictorial presentation of a statement (such as a picture showing how light travels through space). Indicators: "No", "I disagree with…", "They are not right…" Intonation demonstrates disagreement. • An alternative claim is formulated without explicitly contradicting the claim presented. ➢ A statement that contradicts a previous statement is coded as counter-claim even if the previous statement was not coded as claim (because no position-taking was identified). It is then important that the counter-claim explicitly refers to the previous statement and takes a clear position itself (e.g., "That can't be true!")
data (d)	A fact is used as evidence to a claim. (The fact does not necessarily need to be appropriate for the claim from an observer's point of view.) The relation of the fact to the claim must be clear from the wording.	• A fact is expressed (e.g., "That's because snow is white."). • Observations (real situations or mentally constructed scenarios) serve as data (e.g., "Look, there has … happened" or "That is the same as if you would put..."). Also, a phenomenon is provided in order to demonstrate the fact ("Look, when I connect this, the lamp glows") • Events are named which serve as positive consequences to a claim. Negative consequences are listed for a counter-claim. • Students refer to an authority (e.g., "That's what is said in our book.", "That's what our teacher taught us."). • Statements that relate previously expressed data to a claims are also coded as data (e.g., "All mentioned aspects count for …") ➢ For longer utterances, different facts are coded separately as data when not being related to each other. ➢ Justifications such as "because it is this way" or "that's quite clear" are not counted as evidences and are therefore not coded.

warrant (w-i, w)	A coherent statement includes a claim, data, <u>and</u> additional information that describe why/how the data support the claim or relate to the claim.[1] Coding distinguishes between *explicit warrants* (claim, data, and additional information are explicitly presented) and *implicit warrants* (either claim or data are not explicitly mentioned but seem to be included).	*a) explicit warrant (w):* • The statement includes parts that count for a claim, parts that count for data, and parts that describe how the data relate to the claim. Thus, content included into a warrant must be an extension to (previously mentioned) data and claim. ➢ The statement is coded as a whole as warrant. Included claim and data are not coded in addition. ➢ It is not enough to mention data and claim together with an „because", as this does not include any additional information. In such cases, claim and data are coded separately. [2] ➢ A warrant is <u>not coded</u> if claim and data are formulated at a more general level (but without any additional information). ➢ If a claim included into a warrant is not mentioned beforehand, the intention to persuade needs to be clear from the warrant. ➢ Statements such as "Statement A accounts for statement B…" are coded as warrants if the way how these two statements relate to each other is expressed (that these statements relate to each other serves as claim, each statement as data, and taken together with the "how" all of these constitute the warrant). *b) implicit warrant (w-i):* • The interrelationship of data and claim is expressed, <u>either</u> without explicitly mentioning the claim or the data. ➢ For an implicit warrant, the claim needs to be mentioned explicitly in addition to the implicit warrant. ➢ When coding an implicit warrant, the coder should note the explicit warrant next to his coding.
backing (b)	Previously mentioned data, warrants or rebuttals are supported with additional evidence.	• Additional information, law or theory is used to support the previously mentioned data, warrant or rebuttal. ➢ A backing does not need to refer to a (counter-)claim explicitly. • Additional information is presented that explain why a warrant is appropriate to explaining the interrelationship between data and claim.
rebuttal (r)	An evidence (either data, claim, backing, or rebuttal) is refuted. Rebuttals serve a similar function as counter-claims to claims.	• Data, warrants, backings, or rebuttals are refuted in an unspecific manner (e.g., "That isn't true", „No") or the contrary is formulated (e.g., "The object is *not* …"). ➢ This is a rather "simple" way to rebut evidence. ➢ A rebuttal does not necessarily contradict the claim. • The contradiction to the data, warrant, or backing is expressed <u>and</u> reasons are given why the data, warrant, backing is not appropriate or the contradiction itself is elucidated. ➢ This serves as a more expanded rebuttal. ➢ The rebuttal does not necessarily need to be scientifically appropriate. (A "correct" statement can lead to an "incorrect" rebuttal.)

[1] For Toulmin, warrants serve the following function: "Our task is […] to show that, taking these data as a starting point, the step to the original claim or conclusion is an appropriate and legitimate one" (Toulmin, 2003, p. 91).
[2] In contrary, for Toulmin a warrant exists already if someone says "given data D, one may take it that C" (Toulmin, 2003, p. 91)

qualifier (q)	Conditions are mentioned which limit the validity of the claim. However, the claim is not challenged in principle.	• The validity of the claim is restricted (e.g., "This is only true if …", "It depends on whether …") ➢ Note: It is not yet empirically clear whether or not qualifiers also restrict data, warrants, backings, or rebuttals. So far, few instances have been identified that would account for this coding.
justification (j)	used when no clear coding of either data, warrant, or backing is possible	• This value serves as a category "other" for clear phases of argumentation without a valid opportunity to distinguish between data, warrant, or backing.
not coded	Non-argumentative discussions or other activities	

Note. Dots (•) in the last row refer to where this particular element can be found. Arrows (➢) indicate issues and problems that should be taken care of while coding. This coding schema is a shortened version of the German coding schema without any information on how to proceed while coding.

Everyday Example for a Coding

Turn	*Transcript*		*No.*	*TAP*	*Reference*
1.1	S1:	I am sure that Manchester will win today's football match,	1a	c	
1.2		because they are the first team in the league.	1a	d1	turn 1.1
1.3		Whoever is first in the league has the best players and will, therefore, win any football match.	1a	w1	turns 1.1, 1.2
2.1	S2:	I cannot believe that Manchester will win today,	1b	cc1	turn 1.1
2.2		because two of their best players are currently injured.			
3	S3:	Being first in the league has nothing to do with winning all matches.	1	r1	turn 1.3

Note. "No." indicates the successive number of the particular argumentation in the transcript. Thus, the argumentation appearing first in the transcript would be No. 1, the next No. 2 and so on. Letters indicate the different lines within an argumentation. Thus, the letter b refers to elements associated with the counter-claim(s). If no letter is given, it is not clear to which claim or counter-claim the element refers. "TAP" is used to code the argumentation. Here, each element (except of the claim) is numbered so that we can count how many different data, warrants etc. are employed within one argument. () are used to indicate that a previously mentioned element is mentioned anew within the same argumentation. "Reference" serves as the column indicating how the different elements of one argument relate to each other. A final column (not shown in the example) is used to comment on the coding.