

Why Study the Overlap Between “Ought” and “Is” Anyways? On Empirically Investigating the Conventional Validity of the Pragma-Dialectical Discussion Rules¹

Frank Zenker, University of Lund, Sweden

1. Introduction

This paper forwards the (presumably controversial) thesis that the use-value of empirically studying the conventional validity of the pragma-dialectical discussion rules (van Eemeren & Grootendorst 2004: 190-196) is heuristic. This thesis seems natural (to me), if the consequences of a particular theoretical commitment are appreciated: When treating argumentation that supports a descriptive standpoint with a normative premise (aka. a “value sentence”), pragma-dialecticians incur a commitment on the transition from “ought” to “is.” It amounts to embracing the “naturalistic fallacy” as a discussion move that is never appropriate.

In Section 2.1, the aim, method and main result of the recent empirical investigation by van Eemeren, Garssen and Meuffels (2009) are presented. In section 2.2, their conventional validity is discussed. *Vis à vis* the explanation offered by the study’s authors – or so I admit –, the theory internal purpose of this study remains rather unclear to me. After all, as stressed by the authors, and pointed out in Section 3, the normative content of the Pragma-dialectical theory is neither open to refutation by empirical data, nor to confirmation from such data. Therefore, I claim in Section 4, that the *theoretical* value of this investigation is heuristic. Finally, in Section 5, I critically comment on the tension between the level of measurement and the level at which measurement is reported.

2. Studying Conventional Validity Empirically

2.1 Aim, Method and Main Result

The aim is to determine “if and to what extent the norms that ordinary language users (may be assumed to) apply in judging argumentative discussion moves *correspond* to rules which are part of the ideal model of critical discussion” (van Eemeren, Garssen and Meuffels 2009: v; *italics added*). This means to study the rules’ *intersubjective* validity or – insofar as conventions are understood to normally remain implicit – their

¹ Forthcoming in: Eemeren, F.H. van et al. (2011). Proceedings of the 2010 Conference of the International Society for the Study of Argumentation. Amsterdam: SicSat.

conventional validity (see : van Eemeren and Grootendorst 2004: 56, fn. 35). In contrast, the rules' *problem validity* cannot be studied empirically, but is a matter of expert agreement.

As for the method pursued, four out of the ten pragma-dialectical discussion rules are selected: *Freedom Rule*, *Obligation to Defend Rule*, *Argumentation Scheme Rule*, *Concluding Rule*. Based upon these rules, (two to four turn) mini dialogues are created, the last turn of which – on expert opinion – either is or is not a clearly fallacious discussion move (“multiple message design”). Under variation of domains or contexts (*domestic, political, scientific*), the dialogues are presented to lay arguers – mostly younger students – in the form of questionnaires and under the normal precautions for empirical research (e.g., including filler items, in random order, controlling loadedness and politeness of examples, retesting items from previous studies); a sample size of 50 is typical. See van Eemeren, Garssen and Meuffels (2009: 64f.) for examples. Hample (2010) and Zenker (2010) report further details on the method. Notably:

“The third domain [the scientific discussion] was described as the scientific discussion in which – as was emphasized – it was not so much a matter of persuading others but of resolving a difference of opinion in an acceptable manner: Who is right is more important than with whom one agrees.” (van Eemeren, Garssen and Meuffels 2009: 66).

Participants were asked to rate the reasonableness of each dialogue on a seven point Likert scale (1-7). Thus, for each dialogue and each subject, a *reasonableness judgment value* (RJV) is recorded. The RJVs are averaged – yielding an *averaged reasonableness value* (ARV) – and submitted to measures of statistical significance (e.g., mean, standard deviation, effect size).

One thus operationalizes reasonableness as a graded notion. Consequently, the extent can be quantified to which responses from ordinary arguers are (in)consistent with the normative content of the four discussion rules. The value four (4) being the middle point, one reasons: If *this* rule, which generates *these* discourse fragments, is conventionally valid to some extent, then fallacious fragments receive an $ARV < 4$, while non-fallacious fragments receive an $ARV > 4$. And so one compares whether the RJVs fall within the region predicted by expert agreement.

Applied to four out of ten rules, with the exception of the confrontation and the opening stages (van Eemeren, Garssen and Meuffels 2009: 224), the investigation is *non-exhaustive* in the following sense: In principle, violations of different rules (or of a subset of the same rules, but in a different stage) might lead to different results. Furthermore, the ten rule version is a popularization of the more technical 15 rule set (van Eemeren & Grootendorst 2004: 135-157; Zenker 2007, 2009). Precisely how the 15 and the 10 rule set are related is not clear in detail. So, “four out of ten” or “x out of 15” rules have been studied. For a list of fallacies used, see van Eemeren, Garssen and Meuffels (2009: 223).

Under these reservations, the main result is that “(...) the body of data collected indicate that the norms that ordinary arguers use when judging the reasonableness of discussion contributions correspond to a rather large degree with the pragma-dialectical norms for critical discussion” (van Eemeren, Garssen and Meuffels 2009: 224). This claim is principally based on the size of the effect obtained in comparing the ARVs for fallacious and non-fallacious discourse fragments.

2.2 Conventional Validity

Throughout the development of the pragma-dialectical research program, it has been contended that “[t]he [pragma-dialectical] rules (...) are problem valid because instrumental in the resolution process by creating the possibility to resolve differences of opinion” (van Eemeren, Garssen and Meuffels 2009: 27). More precisely, they are considered instrumental to resolving a difference of opinion, insofar as a violation of any rule is understood as a hindrance to this aim.

A further contention is normative in character: The pragma-dialectical rules *should* be conventionally valid, i.e. agreeable to lay arguers. This means, the rules’ content should not conflict with the norms that lay persons (i.e., those not specifically trained in the theory) can be construed to accept. This norm is regularly traced to Barth & Krabbe (1982: 21-22), sometimes to Crawshay Williams (1957), who seems to have been the initial inspiration.

Should these two works harbor a genuine answer to the question why it is important that the pragma-dialectical rules are conventionally valid, then this answer is hidden well. Similarly, neither van Eemeren, Garssen and Meuffels (2009) nor the comprehensive van Eemeren & Grootendorst (2004) offer a detailed explanation either. In

general, at the relevant places (known to me), it is merely stated *that* the rules should be conventionally valid, not why (e.g., van Eemeren, Garssen and Meuffels 2009: 27).

Perhaps an exception is the somewhat detailed explanation found in a 1988 article. From this, three quotes follow. These suggest that the conventional validity of discussion rules – understood as the acceptability of some norms by lay arguers – arises with insight into the rule’s *pragmatic* rationale. In brief, the quotes are consistent with an interpretation according to which intersubjective acceptance comes about through insight into their problem validity. Also see van Eemeren and Grootendorst (2004: 187).

“We believe that the process [of solving problems with regard to the acceptability of standpoints] derives its reasonableness from a two-part criterion: problem-solving validity and conventional validity (cf. Barth and Krabbe 1982: 21-22). This means that the discussion and argumentation rules which together form the procedure put forward in a dialectical argumentation theory should on the one hand be checked for their adequacy regarding the resolution of disputes, and on the other for their intersubjective acceptability for the discussants. With regard to argumentation this means that soundness should be measured against the degree to which the argumentation can contribute towards the resolution of the dispute [i.e., the degree of *problem validity*], as well as against the degree to which it is acceptable to the discussants who wish to resolve the dispute [i.e., the degree of *conventional validity*].” (van Eemeren & Grootendorst 1988: 280)

Pace stylistic changes (e.g., ‘dispute’ would be replaced by ‘difference of opinion’), this is in line with the 2004 presentation. Further in the same article:

“It may now be possible to make plausible that the rules are such that they merit a certain degree of intersubjective acceptability, which would also lend them some claim to conventional validity. [paragraph] The claim of acceptability which we attribute to these rules is not based in any way on metaphysical necessity, but on their suitability to do the job for which they are intended: the resolution of disputes [i.e., their *problem validity*]. The rules do not derive their acceptability from some external source of personal authority or sacrosanct origin. Their acceptability [i.e., their *conventional validity*] *should rest on* their effectiveness when applied [i.e., their *problem validity*]. Because the rules were developed exactly for

the purpose of resolving disputes, they should in principle be optimally acceptable to those whose first and foremost aim is to resolve a dispute. This means that the *rationale for accepting these dialectical rules as conventionally valid* is, philosophically speaking, pragmatic.” (van Eemeren & Grootendorst 1988: 285; *italics added*)

In particular the last sentence suggests (to me) that understanding the rationale of these rules will lead to their acceptance. This interpretation is consistent with that given in van Eemeren and Grootendorst (2004: 187). Note that the rationale being a pragmatic one is not relevant for providing *some* rationale for acceptance. I take it, understanding this rationale necessarily presupposes having understood/learned the rules. Similarly:

“The speech acts which are most useful to all concerned who share a certain goal, for example to resolve a dispute, possess a form of problem validity which may lead to their claim of conventional, intersubjective validity.” (van Eemeren & Grootendorst 1988: 289, n. 14)

Now, *vis à vis* these (admittedly older) quotes, and in the absence of a more recent and superior explanation, it remains unclear (to me) why the pragma-dialectical rules *should* be conventionally valid *independently* of having being learned. Interesting as it is to learn that a methodology supports the claim that they are (or are not), but why *should* they be, to begin with?

If they *are* problem-valid (i.e., acceptable as a solution to a problem), then recognizing their problem-validity expectably brings about their acceptedness, and brings it about for this reason. See van Eemeren and Grootendorst (2004: 187). So, the rules’ problem validity and my (cognitive) ability to recognize their conventional validity – in brief, my appreciating the pragmatic rationale – are sufficient for acceptance (thus, for conventional validity) to come about. If so, then is being acceptable by those *not* trained in these rules somehow important for the theory?

It is trivial to state that the pragma-dialectical (or some other set of problem valid) rules cannot be effective in leading to resolutions, unless at least two persons *de facto* accept them (explicitly or implicitly). So, being conventionally valid clearly is important, because being conventionally invalid would after all mean that – problem valid

rules being unaccepted –, *ceteris paribus*, lay persons may not be expected to maintain a discussion (and obtain a result) that squares with these rules. But why demand (“should”) conventional valid independently of rule acquaintance? Together with the implications on the status of the empirical data, discussed in the next section, the question seems open.

Here, I discount an otherwise important comment by Lotte van Poppel (personal communication). As she points out, it would be less probable for the social aim behind the pragma-dialectical research program of improving argumentative praxis to come about, if the theory’s normative content turned out to be *not* accepted by lay arguers. This, however, cannot merely relate to the exact formulation of this content, and thus must be more than a matter of style. After all, if style did matter, then why investigate conventional validity *in an indirect way*, rather than display the rule set and ask for assent? On this indirectness, see van Eemeren, Garssen and Meuffels (2009: 49f.).

Insofar as the comment then concerns the content (and not various ways of formulating it, e.g., by avoiding/using technical terms), *if* lay arguers and expert judgment do not converge on the content of a set of problem valid rules, well – harsh as this may sound –, so be it! In brief, it is sufficiently unclear (to me) why one assesses (on some methodologically hardened measure) the distance between expert and a lay person judgments. Granted experts find the normative content problem-valid, what kind of support does this content receive from convergence with lay person judgment. Conversely, what kind of doubt arises from divergence?

At this point, it does not help to learn that empirical data take on a special role. As the next section shows, *any* distance between expert and lay person judgment seems to be as theoretically irrelevant as any other.

3. The Special Status of the Results

3.1 Compare, not Test

Compared to applying an *empirical* theory, the data obtained are special: “Empirical data can neither be used as a ‘means for falsification’ nor as ‘proof’ of the problem validity of the discussion rules” (van Eemeren, Garssen and Meuffels 2009: 27). This perhaps deserves explanation. Standardly, an *empirical* theory is tested against experience by applying it to a phenomenon (for which the theory is expected to account), so as to derive a prediction. In the case at hand, the prediction is an expert judgment concerning some discourse item being fallacious or not.

With *A* for *antecedent*, *T* for *theory* and *P* for *prediction*, applying an empirical theory takes the logically valid form: $A; T; (A \& T) \rightarrow P; \text{ergo } P$ (*modus ponens*). If the prediction, *P*, is born out and *A* is not under doubt otherwise, then *T* counts as confirmed. Note that, on a deductive construal, theory confirmation instantiates the invalid schema of *affirming the consequent*. If the prediction is *not* born out (i.e., *non P* is true) and *A* is otherwise not under doubt, then – again, on a deductive construal – falsification instantiates a valid form (*modus tollens*). Treated in deductive logic, only the falsity of ($A \& T$) follows; to derive *non T*, *A* must in addition be less retractable than *T* (see Lakatos 1978; Zenker 2009).

In contrast, the normative content of the pragma-dialectical theory is not tested against lay person judgments, but compared to them (van Eemeren, Garssen and Meuffels 2009: 27). So, some discourse fragment, *A*, to which the pragma-dialectical theory, *T*, is applied does very well deductively imply a prediction, *P*. This is captured by $(A \& T) \rightarrow P$. Yet, *P* and the lay person judgment con- or diverging is without effect on the theory. The explanation offered in defense of this odd support behavior (compared to *empirical* theories) builds on the pragma-dialectical theory offering *norms* rather than *descriptions*.

3.2 Normative vs. Descriptive Statements

The standpoint in van Eemeren, Garssen and Meuffels (2009) is: What lay persons do or do not accept can neither be turned against the theory in the sense of falsification, nor support it in the sense of verification (which, at any rate, requires a non-deductive, viz., an *inductive* construal). Given the comparatively extensive method chapter, the explanation for this standpoint is strikingly brief.

“The presumption in all our empirical studies is that the discussion rules involved are problem valid; the focus is on their conventional validity. The status of the results of this empirical work is special: The empirical data can neither be used as ‘means of falsification’ nor as ‘proof’ of the problem validity of the pragma-dialectical discussion rules. In the event that the empirical studies indicate that ordinary language users subscribe to the discussion rules, it cannot be deduced that the rules are therefore instrumental. The reverse is also true: If the respondents in our studies prove to apply norms that diverge from the pragma-dialectical discussion rules, it cannot be deduced that the theory is wrong. Anyone who refuses to

recognize this is guilty of committing the naturalistic fallacy, the fallacy that occurs when one inductively jumps from “is” to “ought.” (van Eemeren, Garssen and Meuffels 2009: 27)

One may view this quote as a *meta level* assertion, because a claim is raised about the inferential relation between a set of normative statements and a descriptive statement. In effect, the standpoint is that there is no inferential relation at all. This standpoint also shows at object level when evaluating discourse items by which a speaker seeks to support a descriptive standpoint using value statements (normative premises).

“The combination of a descriptive standpoint and a normative argument always leads to an inapplicable argument scheme: The acceptability of a descriptive standpoint is after all independent of the values that are attached to the consequences of the acceptance of that outcome” (van Eemeren, Garssen and Meuffels 2009: 172).

Put generally, “(...) whether something is true or not in a material sense does not depend on the question if we like it or not” (van Eemeren, Garssen and Meuffels 2009: 172). This means, truths (“facts”) do not receive support from, nor can they be undermined by human (dis-)approval.

Pragma-dialectics, of course, is a normative theory. The discussion rules are claimed to be supported by achieving the theoretical value of *problem validity*. This value is achieved through systematically identifying fallacies as hindrances to a resolution oriented discourse. Clearly, to claim problem validity of a normative theory is *not* to assert a norm, but a fact – if it is one. So, lay arguers endorsing norms (in)compatible with the pragma-dialectical ones does not (without committing a naturalistic fallacy) license a claim about the theory’s problem validity.

Just as *undermining norms by facts* is considered fallacious, *supporting facts with norms* is considered equally fallacious. The brevity of stating these contentions indicates that the naturalistic fallacy is a theoretical commitment for pragma-dialecticians. This might surprise. After all, it has long since been recognized that “fallaciousness” depends on various conditions in such a way that “fallacies” may have sound instances. Pragma-dialecticians, however, seem to be committed to the claim that this is not so in the case of the naturalistic fallacy. No further explanation is given.

3.3 The Theoretical Value of Inconsistency

The above boils down to the following: facts (here: the reasonableness judgments of ordinary speakers) are impotent with respect to norms (here: the pragma-dialectical rules). Given this contention, one might wonder why the conventional validity of the pragma-dialectical rules is under study to begin with. After all, even in case the rules *were* conventionally valid – and the claim is that they *are to a rather large extent* – one may at most conclude a conditional claim: *If* ordinary speakers accept rules at all, then these are *not inconsistent* with the normative content of pragma-dialectics.

“Just as would be the case in corpus research, in our series of experiments the conventional validity of the pragma-dialectical rules is investigated not in a direct, but in an indirect sense. Due to the fact that discussion fragments that contain a fallacy are found to be unreasonable by normal judges, and fragments that do not contain any fallacies are deemed reasonable, we deduce that in the judgment of the fairness of argumentation the respondents concerned appeal, whether implicitly or explicitly, to norms that are *compatible, or at least not contradictory*, to rules formulated in the pragma-dialectical argumentation theory”. (van Eemeren, Garssen and Meuffels 2009: 49, *italics added*)

This indirectness comes about for the (above discussed) reason that, by the authors standards, a normative theory cannot be falsified by descriptive data, nor its problem validity confirmed by such data. Hence, *consistency* between the theory’s normative content and the content speakers may be construed to rely on is *rather useless*. *Inconsistency* between the theory and a lay-person judgment has no bearing on the theory either, but *heuristic value*. Simply put, finding inconsistency informs researchers about “what works” without special training and what does not.

4. Heuristics

A number of interesting “anomalies” forthcoming in this study should prove relevant for future theoretical development. Perhaps most importantly, context not only matters, but indeed counts. For example, participants judge an *ad hominem* fallacy to be as reasonable in a *domestic* as in a *political* context, but less reasonable than in a *scientific* context. Similarly, a direct personal attack in a scientific context is judged to be less reasonable

than a *tu quoque* in the same context (ARV = 2.57, standard deviation 0.81 and ARV 3.66, SD 0.86, respectively). Normatively, that the reasonableness value should be the same in all three contexts, and for both variants of the *ad hominem* in the same context, is a defensible claim. Note that nothing in the standard Pragma-dialectical theory explains this dependency on context.

Further, when a standpoint enjoying presumptive status is supported in a fallacious manner, then participants tend to judge this move more leniently than when no such presumption is enjoyed. Normatively again, this makes little sense. Moreover, there are (perhaps striking) differences in culture: some robust effects break down when items are “transported” from one to the next.

Moreover, without training, lay persons will normally not be able to reliably distinguish between a sound *ad absurdum* and a fallacious *ad consequentiam* argument. On the other hand, participants do reliably distinguish the legal principle according to which a presumption of innocence holds unless proven otherwise is, suggesting that further legal principles may generate robust effects as well.

Finally, and perhaps most interestingly, the “trickiness” of the mini dialogues may be varied in future work, to investigate if there is a point at which a variation of content starts to produce effects. After all, the discourse fragments used in this study are conspicuously simple-minded. Some “tweaking” towards more realistic content may lead to certain rules “breaking down,” for also this study supports the claim that participants tend to be influenced by the content of a standpoint: If you assent to what is supported by fallacious means, you will judge such fallacies more leniently than you would, if you did not assent. While perhaps understandable, this makes little normative sense.

5. Data Reporting

Before concluding, a critical point pertains to the tension between the level of measurement and the level at which measurement is reported. As pointed out above, measurement occurs on a seven point Likert scale: 1, 2, 3, 4, 5, 6, 7 (see Likert 1932). This means, reasonableness judgments are recorded at *ordinal* level. By definition, at this level, we lack information on distances between points. So, it is for example *not known* if the distance between the points 5 and 6 is the same as that between 2 and 3.

When reporting and statistically treating data, the implicit assumption is that the distance between any two scale points is *the same*. This assumption is needed, for oth-

erwise averaging – which yields fractions (e.g., an averaged reasonableness value of: $2\frac{200}{375}$) – would be meaningless. Thus, data are treated as if they had been measured at *interval* level. Though deeply entrenched, this assumption is open to doubt. Treating it should make for an interesting case study on a long-lasting scientific controversy. See Jamieson (2004) and Norman (2010) for both positions.

The standard report formats are the *mean plus standard deviation*. The mean is the sum of all measurement-values divided by the number of measurements. To indicate the spread of data are around the mean, the *standard deviation*, s , is used (where x is a data value, \bar{x} the mean, and n the number of measurements).

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}$$

The standard deviation is a widely accepted measure of dispersion. However, the value of s will not allow reconstructing the *exact* spread. Thus, until they do their own experiments, readers remain ignorant as to how many subjects showed what deviation in their reasonableness judgments. This, unfortunately, makes the data less useful for replication than one might desire. In fact, by how much *individuals* deviated in their judgments is perfectly hidden.

This suggests that the aim of the entire empirical study was not to find out and then report *precisely* which reasonable values subjects assigned to an item. Rather, the point was to show that, for the mini dialogues constructed (the great majority of which suffer from a near-triviality of content), theoretical prediction and averaged lay person judgment converge. In other terms, the results strongly suggest that it is possible to construct examples which lay persons can distinguish – on average and to a rather large extent – into fallacious and non-fallacious arguments.

6. Conclusion

As pointed out, the theoretical point of comparing expert and lay person judgments concerning the reasonableness of rule-generated discourse fragments remains to be explicated. In the absence thereof, the naturalistic fallacy may count as a theoretical commitment of pragma-dialecticians. Whether this commitment needs additional justification depends on one's prior theoretical commitments. At any rate, it should be interesting to ask a pragma-dialectician for such a justification.

Several examples of the heuristic value of the empirical investigation of the conventional validity of four of ten pragma-dialectical discussion rules were pointed out. On pains of having appearing too critical in this text, readers are reminded of two reviews (Hample 2010, Zenker 2010). These agree in praising van Eemeren, Garssen and Meuffels (2009). Notwithstanding the criticism raised here, the study is highly relevant, irrespective of one's theoretical background.

7. References

- Barth, E.M. and E.C.W. Krabbe (1982). *From Axiom to Dialogue: A Philosophical Study of Logics and Argumentation*. Berlin: De Gruyter, 21-22.
- Crawshaw-Williams, R. (1957). *Methods and Criteria of Reasoning: An Inquiry into the Structure of Controversy*. London: Routledge and Kegan Paul.
- Eemeren, F.H. van and R. Grootendorst (1988). Rationale for a Pragma-Dialectical Perspective. *Argumentation* 2, 271-291.
- Eemeren, F.H. van and R. Grootendorst (2004). *A Systematic Theory of Argumentation. The Pragma-dialectical approach*. Cambridge: Cambridge University Press.
- Eemeren, F.H. van, Bart Garssen and Bert Meuffels (2009). *Fallacies and Judgments of Reasonableness. Empirical Research Concerning the Pragma-Dialectical Discussion Rules*. Dordrecht: Springer.
- Hample, D. (2010). Review of Fallacies and Judgments of Reasonableness. *Argumentation (online first)*.
- Jamieson, S. (2004) Likert scales: how to (ab)use them. *Medical Education* 38, 1212-1218.
- Lakatos, I. (1978). *The Methodology of Scientific Research Programs*. Cambridge: Cambridge University Press.
- Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology* 140, 1-55.
- Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Science Education (online first)*.
- Zenker, F. (2007). Pragma-Dialectic's Necessary Conditions for a Critical Discussion. In Hansen, H. et al. (eds). *Proceedings of the 7th Int. Conference of the Ontario Society for the Study of Argumentation (OSSA)*, Windsor, ON, CD ROM.
- Zenker, F. (2009). *Ceteris Paribus in Conservative Belief Revision*. Frankfurt: Peter Lang.
- Zenker, F. (2010). Review of Eemeren, F.H. van, B. Garssen & B. Meuffels (2009). *Fallacies and Judgments of Reasonableness*. Amsterdam: Springer. *Cogency* 3 (forthcoming).