sciendo

STUDIES IN LOGIC, GRAMMAR AND RHETORIC 68 (81) 2023

DOI: 10.2478/slgr-2023-0007

Frank Zenker
College of Philosophy, Nankai University, Tianjin, P.R. China
e-mail: frank.zenker@nankai.edu.cn
ORCID: 0000-0001-7173-7964

Kamila Dębowska-Kozłowska
Faculty of English, Department of Pragmatics of English, Adam Mickiewicz University, Poznań, Poland
e-mail: kamila@amu.edu.pl
ORCID: 0000-0002-5039-6133

David Godden
Philosophy Department, Michigan State University, East Lansing, Michigan, USA
e-mail: dgodden@msu.edu
ORCID: 0000-0002-4870-6331

Marcin Selinger
Department of Logic and Methodology of Sciences, University of Wrocław, Wrocław, Poland
e-mail: marcin.selinger@uwr.edu.pl
ORCID: 0000-0002-9685-6684

Simon Wells
School of Computing, Edinburgh Napier University, Edinburgh, Scotland
e-mail: s.wells@napier.ac.uk
ORCID: 0000-0003-4512-7868

THOU SHALT NOT SQUANDER LIFE – COMPARING FIVE APPROACHES TO ARGUMENT STRENGTH

Abstract. Different approaches analyze the strength of a natural language argument in different ways. This paper contrasts the *dialectical*, *structural*, *probabilistic* (or Bayesian), *computational*, and *empirical* approaches by exemplarily applying them to a single argumentative text (*Epicureans on Squandering Life*; Aikin & Talisse, 2019). Rather than pitching these approaches against one another, our main goal is to show the room for fruitful interaction. Our focus is on a dialectical analysis of the *squandering argument* as an argumentative response that voids an interlocutor's right to assertion. This analysis addresses the pragmatic dimensions of arguing and implies an argument structure that is consistent with empirical evidence of perceived argument strength. Results show that the squandering argument can be evaluated as a (non-fallacious) *ad hominem* argument, which however is not necessarily stronger than possible arguments attacking it.

Keywords: argument structure, Bayesian, computation, diagram, dialectic, empirical, evaluation, perceived argument strength, thought listing.

1. Introduction

A central concept in the study of natural language argumentation, *argument strength*, can be studied using a range of theoretical approaches. To appreciate their differences, we contrast the *dialectical*, *structural*, *probabilistic* (or *Bayesian*), *computational*, and *empirical* approaches. Sequences of dialogue moves (attacks and counterattacks) are abstracted away in all but the dialectical approach. Dialectical insights can nevertheless be incorporated into the structural argument models of other approaches. So, rather than having to pitch the five approaches against one another, we seek to examine the room for fruitful interaction. Indeed, all five approaches contribute to, and inform, the final analysis.

Within each approach, analysts first reconstruct the argumentative structure and then identify the strongest argument. Each approach thus is computational because analysts must compute an *order* among arguments corresponding to their comparative strength. To arrive at this order, the empirical approach leverages the cognitive responses of experimental participants. The Bayesian approach uses a probabilistic semantics, whereas the structural and the computational approaches can use different semantics. For the dialectical approach, finally, computing the order *is* to diagram its argumentative structure.

These five approaches are here exemplarily applied to the main arguments in *Epicureans on Squandering Life* (Aikin & Talisse, 2019; see Aikin & Talisse, 2022). What proves crucial to the analysis is the status of the *squandering argument* as an argumentative response that voids an interlocutor's right to assertion. On a dialectical analysis, which addresses the pragmatic dimension of arguing, the argument's response-status implies a structure with *dependent* lines of support. This structure proves to be consistent with indirect empirical evidence of perceived argument strength. The squandering argument can thus be evaluated as a strong (non-fallacious) *ad hominem* argument, although it is not necessarily stronger than possible arguments attacking it.

We begin by presenting the main arguments in Aikin and Talisse (2019) (Sect. 2), give a brief overview of the five approaches (Sect. 3), and then contrast how each approach evaluates argument strength in application to the main arguments in *Epicureans on Squandering Life* (Sects. 4.1–4.5). We conclude in Sect. 5.

2. Epicureans on Squandering Life

The main arguments in *Epicureans on Squandering Life* (Aikin & Talisse, 2019) put forward in support of the conclusion (S) 'we should not fear death' are threefold: the *no subject of harm argument* (NSH-Arg), the *symmetry argument* (SYM-Arg), and the *squandering argument* (SQU-Arg).

As an analysis of natural language argumentative text comes with degrees of freedom, analysts must exercise choices about the resolution at which arguments are diagrammed and the implicit premises that are made explicit. A crucial choice for Aikin and Talisse's (2019) text is whether the SQU-Arg is analyzed as an *independent* or a *dependent* line of support for the conclusion S. We argue that the NSH-Arg and SYM-Arg provide independent lines of support, whereas the SQU-Arg is an example of "developmental pieces of reasoning [...] presented in the thick of an exchange between particular discussants" (Aikin & Talisse, 2019). Because the SQU-Arg defends the conclusion (S) 'we should not fear death' by *responding* to the premise 'life is a valuable thing (for us)', it is best analyzed as a *dependent* line of support for S.

The premises of the NSH-Arg are (1) 'there is no subject to feel pain once we are dead' and (2) 'pain is the only truly bad thing'. The SYM-Arg also relies on two premises: (1) 'the time before we were born is relevantly similar to the time after we die' and (2) 'we do not feel dread with respect to the time before our birth.' The NSH-Arg and SYM-Arg both support the intermediate conclusion (RF) 'there is no reason to dread the time after our death,' and its modal variant (S) 'death *should* not be feared' or simply 'we *should* not fear death'.

The SQU-Arg, by contrast, responds to the *death-as-deprivation* counter-argument (DD-Arg) that *there is* a reason to fear death, namely: 'death deprives us of something valuable', i.e., our life. The SQU-Arg attacks the DD-Arg on the grounds that the DD-Arg's proponent *fails* to treat their life as valuable by not appreciating the simple joys that life offers, e.g., "savour[ing] the warmth of the sun or the cool of the evening" (Aikin & Talisse, 2019). Instead, the proponent is said to *waste* their life by pursuing things that never please them, making their life pathetic (e.g., by striving for better careers, more money, being addicted to political and mass-mediated revelations, or observing a strict diet, etc.). This is illustrated in Fig. 1, where the sequence of dialogue moves is omitted and only the argument structure is presented.

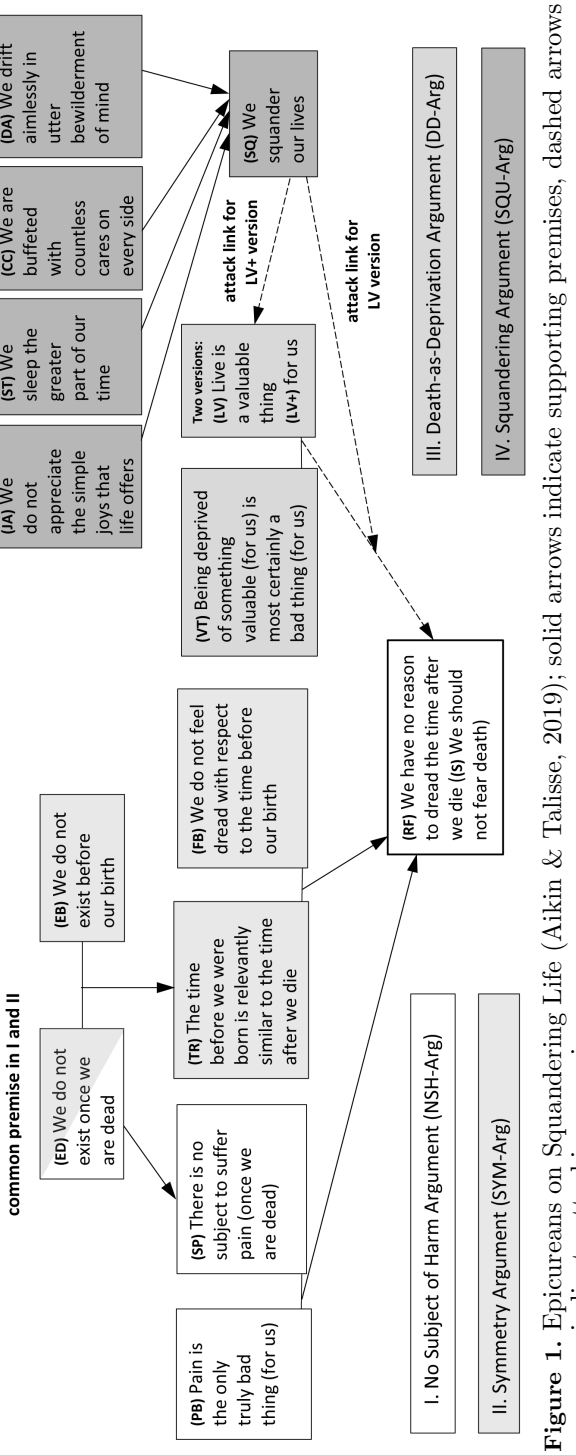

Figure 1. Epicureans on Squandering Life (Aikin & Talisse, 2019); solid arrows indicate supporting premises, dashed arrows indicate attacking premises

3. Overview

A *dialectical approach* models argument strength as discussants' commitments, entitlements, and obligations at some dialogue stage. These collectively comprise the available *move space*. Argument strength is operationalized as the (un)availability of discussant moves that constrain future discussant moves, which is minimally a function of non-losing future moves. Given the move space, the strongest proponent argument leaves no available opponent move, and the weakest proponent argument constrains no opponent move. Because analysts must discover the move space heuristically – there is no infallible algorithm – the strongest *possible* argument may differ from the strongest *actual* argument an analyst identifies.

The same holds for a *structural approach*, the objects of which are separate inferences constituting atomic simple or linked arguments. These objects form complex arguments via syntactic operations corresponding to convergent, divergent, and serial arguments (standard part), as enriched with counter-considerations and undercutters (dialectical extension). To evaluate arguments, analysts combine the values of basic premises with the weights of component inferences in an order corresponding to the entire structure. Suitable operations on values and weights then extend the evaluation function's domain to obtain the value of the conclusion, representing an argument's strength. What a suitable operation *is*, however, remains undefined.

Provided non-zero prior probability values for reasons and claims, a *Bayesian approach* models argument strength by deriving a claim's posterior probability from its likelihood. Argument strength refers to the numerical difference (as measured) that the credences one associates with specific reasons make to the credences one associates with specific claims. The all-things-considered *best* formal measure of argument strength is a matter of debate. The subjective bent of the Bayesian approach shows most clearly if the content of a reason-claim complex rests on single-event probabilities. These cannot be synced with any objective frequency other than the number of those who adhere to this content. But as an evaluative criterion, this invites circularity.

Though rarely referring explicitly to argument strength, a *computational approach* makes room for three concepts of argumentative strength: (i) based on "surviving" arguments after the application of a semantics, (ii) based on comparing different semantics, and (iii) based on preference orderings. The first two concepts are objective, scalable, algorithmic, and generally applicable. But their high granularity can lead to possibly many, equivalent sets of "strong" arguments. The third concept, by contrast, corre-

sponds to a subjective, effort-intensive, and more specific sense of argument strength, a sense that can be fine-tuned to arguers' preferences.

Finally, on an *empirical approach*, analysts can study argument strength either directly by relying on a prior notion of argument strength (e.g., a measure of evidence quality) or alternatively, argument strength can be studied indirectly based on a population's cognitive responses. Since an empirical approach is data-driven, it can be tailored to contextual factors. The main challenge is to identify the specific set of indicators that provide the most descriptively adequate model of perceived argument strength.

Let this suffice as an overview and now turn to the details.

4. Comparing five approaches to argument strength

4.1. The dialectical approach

4.1.1. Argument strength dialectically conceived

A dialectical approach involves the analytical task of reconstructing written or spoken discourse as turn-by-turn sequences of dialogue moves, or speech acts, followed by evaluating them using dialogue game-protocols or -rules. Argumentative norms are modeled as *procedural* rules (aka protocols) that permit, oblige, or prohibit particular moves or move sequences. This creates obligations or permissions to make moves of specific kinds at future dialogue stages. For example, if a discussant has the right to assert any standpoint, then a proponent having asserted a standpoint P is obliged to defend P if challenged, thus acquiring a burden-of-proof for P. And the respondent has the right to challenge any asserted standpoint.

Modeled as the set of commitments, entitlements, and obligations that pertain to discussants at a given dialogue stage, argument strength can be understood as the *limited availability* of participant moves that constrain further interlocutor moves. These moves collectively comprise the available move space. The strongest proponent argument, itself a function of available non-losing future participant moves, leaves no further opponent move except concession (i.e., retraction of standpoint or critical doubt). And the weakest proponent argument given the move space constrains no opponent move.

To determine the move space, *informal* dialectical theories (e.g., van Eemeren & Grootendorst, 2004; Walton, Reed & Macagno, 2008) draw on conceptual, analytical, and evaluative tools that include argumentation schemes, critical questions, and fallacies. When arguers give reasons, they invariably draw upon a repertoire of argument schemes, the acceptability

of which varies with a discursive domain or field. In activating an argument scheme, available dialogue moves draw upon both substantial premises and rules. Unavailable moves, by contrast, which against background information and context are criticizable as fallacies, constitute the negative move space. This space is delimited by the critical questions associated with each scheme, questions that test for typical ways of using a scheme instance infelicitously.

Such informal tools track moves on a turn-by-turn basis, helping to determine how the burden-of-proof (BOP) shifts according to the obligations that each discussant incurs. At the end of a dialogue, a proponent who fails to discharge an initial BOP must retract their standpoint as insufficiently supported. Whereas if a proponent meets their BOP by offering a dialectically strong enough argument, the respondent must withdraw their critical doubt. Informal tools generally have heuristic value. But none delivers a comprehensive catalog of (im-)permissible arguments, nor an exhaustive list of (im-)permissible discussant moves, let alone a complete list of criteria for cogent arguments.

A more systematic approach is found in *formal* dialectical theories (Barth & Krabbe, 1982; Hamblin, 1970, 1971; Kieff, 2011; Krabbe, 2013; 2017; Krabbe & Walton, 2011; Rescher, 1977; Walton & Krabbe, 1995). These theories depict arguments as *dialogue profiles* that are distinguished by different protocols (Krabbe, 1999; 2002; van Eemeren et al., 2014, 366–367; Walton, 1999, 54f.; 2015, 96f.). Formally, a dialogue profile is a directed graph with a tree structure. Tree-nodes represent *possible* dialogue moves at a dialogue stage given a dialogue's rules, and edges (joining the nodes) represent paths to *permissible* moves.

Any actual dialogue thus instantiates a tree branch – a particular path from root to tip. An obligatory move corresponds to a single path from a previous move, whereas a permissible move gives rise to several different paths. If a graph's edges are weighted equally, *path connection strength* constrains the available response moves, provided the strategic goal is to execute move sequences that compel interlocutors to one of two game-ending moves: proponent-standpoint retraction or respondent-standpoint concession. A discussant thus has a *winning strategy* if, and only if, an available sequence of moves lets each subsequent branch terminate in a losing interlocutor move.

4.1.2. Position strength

Argument strength as dialectically conceived thus is a function of the (un)availability of permissible move sequences originating at the present di-

alogue stage and ending in a discussant's role-specified goal being achieved. Given that a respondent's *primary* dialectical role is to raise critical doubt about a standpoint (rather than to defend a standpoint under doubt), a respondent fails to offer *standpoint-supporting* reasons unless a counter argument or an alternative standpoint are advanced to motivate critical doubt. Therefore, if argument strength narrowly refers to the (un)availability of (non-)losing future moves, it can also be labeled *position strength*.

This seemingly clear evaluative standard, however, is complicated by situational and structural considerations that pertain to *background commitments*, to *changes* in the commitment set, and to *meta-argumentation*, to which we now turn.

First, determining the availability of a discussant's non-conceding moves requires knowledge of the discussant's background commitments. Proponents must be thought to strategically select premises they expect to be most resistant to critical doubt or even to go unchallenged. Similarly, respondents must be thought to strategically attack moves they expect to be most vulnerable or even indefensible. Because one cannot assume that unchallenged discourse material is automatically accepted, proponent commitments relevant to evaluating position strength may remain in the background. A discussant *not* making a specific move nevertheless indicates a disinclination to making it. For instance, if not voicing an objection indicates having no objection, then letting discourse material pass unchallenged commits to it.

Second, if "winning" an argument depends on the (un)availability of (non-)losing future moves, then some moves are better positioned than others. Skeptics, for instance, can dogmatically maintain doubt in view of otherwise unobjectionable proponent claims. Similarly, proponents can constrain their interlocutor's move space by contingently accepting implausible commitments or denying plausible ones. Especially if commitment retraction is permitted, an evaluation of argument strength should therefore consider whether a discussant's opening and closing commitment sets are (in-)coherent, for instance by using a measure of *minimum mutilation* (Quine, 1961, 44; 1992, 14–16).

Third, the praxis of giving reasons comprises the meta-argumentative task of evaluating reasons, as well as the meta-dialogical task of critiquing evaluative standards (Finocchiaro, 2007; 2013; Krabbe, 2003). The move space, therefore, must allow room for both tasks. A fallacy accusation, for instance, is a meta-argumentative move that to defend against may involve critiquing an applicable evaluative standard. A second example is to evaluate not a speech act's content, but a discussant's *entitlement* to perform the act. A case in point is the following dialectical analysis of the SQU-Arg.

4.1.3. A dialectical analysis of the squandering argument

A dialectical analysis of the SQU-Arg considers how the respondent (RES) replies to the proponent (PRO) with a charge of *pragmatic inconsistency*. PRO raises the NSH-Arg and SYM-Arg to support the standpoint (RF) 'death is not a bad thing to be feared'. RF provides the grounds for the modal claim (S') 'the living *should* not fear death', respectively (S) 'we should not fear death'. On this analysis, the features of the NSH-Arg and SYM-Arg that account for their strength are nothing other than the discussants' commitment to the arguments' premises and inferences.

The premises of the NSH-Arg are: (1) 'after death, there is no experiencing subject for pain', because (1.1) 'the dead do not constitute experiencing subjects'; and (2) 'pain is the only truly bad thing to be feared'. Given semantic knowledge, it follows from (1) that (3) 'the dead cannot experience anything – including pain', wherefore (RF) 'death is not a bad thing to be feared'. The SYM-Arg rests on the premises that (4) 'the time after death is relevantly similar to the time before birth', and that (5) 'we do not dread the time before birth', wherefore – given that 'we should not fear the time after death' – it follows from 4 and 5 by analogy that (RF) 'death is not a bad thing to be feared.'

RES's reply, the DD-Arg, challenges neither the premises nor the inferences of the NSH-Arg or the SYM-Arg. The DD-Arg rather is a counter-argument, the conclusion of which is (S*) 'there is a reason to fear death'. Since the DD-Arg uses the NSH-Arg's premise 1.1, this secures its dialectical acceptability. (See the shared premise in Fig. 1). With the DD-Arg, RES relies on the premises that (6) 'experiencing life is a good thing', wherefore, given 1.1, (7) 'death deprives of the good things that life is'. Granted that (8) 'being deprived of a good thing is bad', RES thus arrives at (S*) 'death is a bad thing to be feared'.

As for the dialectical strength of the case for S, once RES is a proponent for S* (entailing not-S), the dialogue is *mixed*: each discussant maintains a distinct standpoint (S, S*) and each incurs dialectical obligations to defend them. Since the dialogue ends with PRO contending that RES should retract the DD-Arg, while PRO's NSH-Arg and SYM-Arg are unchallenged, the argument has a *conductive* structure. Evaluating the support for S and S* thus requires a weighing of competing pro- and con-reasons.

To see this, consider that PRO responds to the DD-Arg by challenging (7) 'death deprives of the good thing that life is.' This maneuver entails that RES takes 7 to follow from 1.1 and 6, neither of which PRO denies. PRO in turn challenges RES's inference from 1.1 and 6 to 7, claiming that not-7 follows from 1.1 and (8) 'one is deprived of something only if one exists'.

From the contraposition of 8 (non-existence precludes possible deprivation) and 1.1, it then follows by *modus ponens* that (9) 'the dead are not deprived of anything (including the good that life is)'. Thus, 9 entails not-7.

RES's counter-response, the *Less of a Good Thing Argument* (LGT-Arg), hypothetically grants (9) – i.e., the dead are not deprived of anything – but claims that (11) 'by dying, the living lose something good, namely the experience of life', would follow from (10) 'the experiences of life are good' and (1.1) 'the dead do not exist to experience anything'. Hence, (S*) 'death is a bad thing to be feared," entailing not-S.

As the dialogue now leads to the SQU-Arg, given that PRO has not denied 10 – 'the experiences of life are good' – PRO would by accepting 1.1 seem committed to 11, and hence to not-S. But rather than concede by withdrawing S, PRO raises the following *ad hominem* (i.e., the SQU-Arg): 'You say you fear death because it deprives you of the life you value, but *look at what you do with that life*!' Rather than denying 10, PRO thus denies that RES is *entitled* to assert 10, because discussants like RES deny themselves life's simple pleasures, pursue insatiable desires, fear what one has no reason to fear (e.g., death), and so live with self-inflicted torment, *squandering* life.

PRO's claim thus is that (12) 'one who truly feared death would not squander life.' Although PRO accepts 10 – 'the experiences of life are good' – PRO specifically claims that those who squander life demonstrate a failure to value life (i.e., do not display the right commitment to life), wherefore RES is *unentitled* to assert 10. With the SQU-Arg, therefore, PRO objects that RES's contending 10 is pragmatically inconsistent.

4.1.4. Evaluation

That PRO and RES rely on premises the other accepts, yet disagree about the premises' consequences, speaks to what Pragma-dialecticians call an argument's *intersubjective* validity (vs. its objective or extra-discursive validity). As rational critics, PRO and RES offer counter-argument, thus becoming proponents for a standpoint that contradicts their opponent's standpoint. Specifically, once RES forwards the DD-Arg as a potential defeater to PRO's NSH-Arg and SYM-Arg, PRO replies to RES's DD-Arg *not* by requesting a demonstration that 7 follows from 1.1 and 6 (which would place the BOP upon RES), but by seeking to rebut the DD-Arg.

PRO and RES thus take on new discursive commitments and probative burdens which seem to *weaken* each interlocutor's position. Because the greater the number of commitments, the greater is the number of obligations to take on additional commitments in their defense. Since this further restricts the move space and raises the risk of retraction, it makes *additional*

response options available to the opponent, yet without taking on new commitments. Conversely, strengthening a position entails limiting new commitments to those an interlocutor is already committed to, and from which the contested claim follows. Ideally, leveraging an opponent's commitment restricts the respondent's move space to retracting critical doubt (standpoint concession) or, less ideally, to retracting prior commitments the contested claim follows from.

Although RES and PRO manage as much, PRO is committed to every claim of RES's LGT-Arg *except* its conclusion, not-S. Yet, PRO's response to the LGT-Arg does not retract any commitment or standpoint, nor does PRO deny that not-S is entailed, nor is PRO committed to this entailment. This shows that PRO's focal move, the SQU-Arg, is meaningful only in a framework that recognizes *challenges to the act of making* a dialogue move (here: RES's right to assert 10). A framework that represents only the argument content cannot properly capture this move.

Insofar as a dialectical framework shows the alleged *fallaciousness* of RES's LGT-Arg – itself residing in the pragmatic inconsistency of asserting 10 – and so shows the *non-fallaciousness* of PRO's *ad hominem* argument that the SQU-Arg is, the SQU-Arg would be strong if it restricts RES's move space to conceding that people live the sort of life an Epicurean extolls. At this moment, a meta-dialogue would presumably begin. Although its content and pragmatics are unavailable, PRO would need to defend S – 'death is not a thing to be feared' – while committing to 10 – 'the experiences of life are good'. This is a prima facie difficult case to make. The LGT-Arg thus seems strong, outweighing the NSH-Arg, and possibly even the SYM-Arg. Yet RES would simply be where the discourse already stands. That is, RES would need to offer an *alternative* defense for not-S, or else retract not-S while raising doubt regarding PRO's NSH-Arg and SYM-Arg to compel PRO to retract S.

4.2. The structural approach

4.2.1. Argument structure and abstract evaluation

Employing argument diagrams as used in informal logic, the structural approach provides a bottom-up, general, and abstract model to evaluate the strength of arguments, a model that is compatible with various argumentation frameworks (*cf.* Gordon & Walton, 2006, 2015; Prakken, 2010; Modgil & Prakken, 2014; Thomas, 1983; Tokarz, 2006; Selinger, 2014). In the model's *standard part*, the objects under evaluation are separate inferences constituting atomic arguments (i.e., simple or linked structures).

Atomic arguments form complex ones via syntactic operations that result in convergent, divergent, and serial argument structures. The model's *dialectical extension* includes three types of counter-considerations and counter-arguments: an *underminer* attacks premises, a *rebuttal* attacks a conclusion, and an *undercutter* attacks an inferential link.

This generalized model abstracts from the specific values representing argument strength and from the specific algorithms transforming the acceptability of premises into the acceptability of a conclusion. The model assumes that all arguments consist of sentences of a predefined language L. *Argumentation structures* can be represented as finite, non-empty sets of *sequents*, i.e., quadruples of the form $\langle P, c, d, R\rangle$, where $P \subseteq L$ is a finite, non-empty set of premises, $c \in L$ is the conclusion, $d \in \{0, 1\}$ is a Boolean value (1 for pro- and 0 for con-sequents), and $R \subseteq 2^L$ is a finite set whose elements are non-empty, finite sets of (linked) undercutters (Selinger, 2019). This approach to modeling argumentation structures provides an alternative to argument graphs as proposed by Walton and Gordon (2006, 2015) and to Prakken's approach (2010) which combines Vreeswijk's (1993, 1997) formalization with Pollock's (1987) definitions of defeat.

For evaluative purposes, the model assumes two types of values. Values assigned to sentences are simply called *values*; values assigned to inferential links are called *weights*. The set of values is any set V (with at least two elements) whose elements can be assigned to the sentences of a subset $L' \subseteq L$ by a partial evaluation function v. Elements of a distinguished (non-empty) proper subset $V^* \subseteq V$ are assigned to valid sentences, which are usually equated with audience-accepted sentences. Similarly, the set of weights is any set W (with at least two elements) whose elements can be assigned to direct inferences, regardless of the premises' values. Elements corresponding to valid inferences are represented by a proper (non-empty) subset $W^* \subseteq W$.

The sets V and W are ordered according to the 'is stronger than'-relation. This relation determines the corresponding concept of argument strength for a given *argumentation system*. Such a system is defined by (i) the sets of values and weights (V, W), (ii) their distribution over the sentences and inferences of L, and (iii) the operations on them that correspond to the structure of arguments formulated in L. (If unordered, V and W determine no such concept.) This ordering merely assumes that any distinguished value or weight is stronger than any undistinguished one. Prima facie, however, the natural choice seems to be a linear (total) order, which leaves all computable arguments pairwise comparable. The largest element of V, if any, can be interpreted as the full acceptance of a given sentence, and the least element, if any, either as its full rejection or its undecidability.

Similarly, the largest element of W, if any, can be interpreted as deductive (strict) inference, and the least element, if any, either as deductive rejection or irrelevance.

In the 2015 edition of the *Carneades* system, for instance, Walton and Gordon (2015) use the set of values *out*, *undecided*, *in* and the set of weights consisting of the real numbers of the interval $[0, 1]$. Tokarz (2005), by contrast, uses the set of values $\{1, 2, 3, 4, 5\}$ and the scale of weights (*nil*, *weak*, *moderate*, *strong*, *deductive*) as defined by Thomas (1986). But whereas Thomas used the set of values *untrue*, *true*, Tokarz replaces this scale with a five-element scale consisting of the same set of integers as the set of values, i.e., $\{1, 2, 3, 4, 5\}$. By extending Tokarz's scale to the infinite set of rational numbers from the interval $[0, 1]$, Selinger (2014) represents both values and weights. Finally, ASPIC+ (Modgil & Prakken, 2014) can be interpreted as using merely two values, i.e., *accepted*, and *not accepted*. However, the fact that its underlying language is assumedly closed under negation allows us to distinguish *undecided* and *rejected* among *not accepted* sentences. Inferences, on the other hand, can merely be *valid* or *not valid*. (While a strict inference is always valid, a defeasible inference is so, only if it falls under an applicable rule.)

To achieve an evaluation, the values of the first premises (as assigned by an initial *evaluation function*) are combined with the weights of the component inferences in an order corresponding to the structure of the whole argument. Using suitable operations on values and weights, the initial evaluation function's domain is then extended step-by-step to obtain the value of the conclusion. This value is defined as the *strength of the argument*. Counter-arguments can be evaluated as separate arguments or be combined/aggregated with arguments they themselves attack, to thus evaluate the aggregated whole (Selinger, 2019). The *relative strength* of a counter-argument thus is the "gap" between the strength of an attacked argument and that of the aggregated one. (Any such attack succeeds if the value of the aggregate's conclusion is excluded from the distinguished set V^*.)

This evaluation suffices to implement the RSA criteria for good arguments: relevance, sufficiency, and acceptability (Johnson & Blair, 1977). The initial evaluation function corresponds to premise acceptability, the weighing function to premise relevance, and a combination of both values and weights to premise sufficiency. The evaluation also complies with general conditions that any satisfactory formal theory of informal logic should meet (see Walton & Gordon, 2015).

Potentially impeding the evaluation are logical interdependencies between an argument's components. With convergent arguments, for instance,

argument strength may be overestimated (*double counting fallacy*), whereas the acceptability of sets of premises may be underestimated (the complement of the *double counting fallacy*). Also vulnerable to this fallacy is the recalculation of already assigned values (e.g., to the first premises that have been undermined). If these interdependencies are overlooked, their effect on argument strength remains unaccounted for. Moreover, analysts must assess each argumentation structure as to whether a rebuttal rebuts only its conclusion or also undercuts an opposite pro- or con-argument that (respectively) supports or rebuts the conclusion.

We now apply a simple, specified instance of this generalized model to the SQU-Arg, here starting from the definitions of the sets of values and weights, and operations on them.

4.2.2. A structural analysis of the squandering argument

Let the set of values V be $\{0, \frac{1}{2}, 1\}$ and let its distinguished subset V^* be $\{1\}$, where 1 stands for 'acceptable', 0 for 'rejected', and $\frac{1}{2}$ for 'undecided'. Further, let the set of weights W be $\{0, 1\}$ and let its distinguished subset W^* be $\{1\}$, where 1 stands for 'valid' and 0 for 'not-valid'. The value of a linked set of sentences is the *minimum* value assigned to any of its elements. The initial weight of any deductive, and (by default) of any defeasible inference is 1, although the latter can be transformed into 0 if a relevant undercutter is accepted. Thus, the value of the conclusion of any pro-sequent, i.e. its strength, is 1 if, and only if (*iff*), the value of the set of its premises is 1, the (initial) weight of its inference is 1, and there is no set of (linked) relevant undercutters of value 1; otherwise, this strength is $\frac{1}{2}$. By parity of reasoning, the strength of any con-sequent is 0, *iff* the value of the set of its premises is 1, the (initial) weight of its inference is 1, and there is no set of (linked) relevant undercutters of value 1; otherwise, this strength is $\frac{1}{2}$.[1]

The strength of a convergent pro-argument is the *maximum* strength of its converging sequents, whereas the strength of a convergent con-argument is the *minimum* value. Since pro-arguments normally cannot take value 0, nor can con-arguments normally take value 1, the strength of a *conductive* argument can be defined as 0, *iff* the overall strength of its con-arguments is 0, as $\frac{1}{2}$ *iff* both pro- and con-arguments take the overall value $\frac{1}{2}$, and as 1 *iff* the overall strength of pro-arguments is 1 and that of con-arguments is $\frac{1}{2}$.

The structure in Fig. 1 consists of four arguments, where solid arrows represent support links and dashed arrows attack links. As pro-arguments in support of (RF) 'we have no reason to fear death', the NSH-Arg

and the SYM-Arg provide the primary (positive) part of the structure. The DD-Arg is a con-argument that rebuts the pro-arguments. And the SQU-Arg undercuts *or* undermines the DD-Arg, depending on the interpretation of one of the latter's premises: LV as 'life is valuable' (the absolute version) *vs.* LV+ as 'life is valuable *for us*' (the relative version). More precisely, since the SQU-Arg denies LV+ but does not deny LV, it undermines the DD-Arg's relative version, while undercutting its absolute version.

Because the scope of quantification affects the logical relations among components, this affects the evaluation. To keep things simple, in Fig. 1 the scope of quantification of all premises, as well as of the conclusion RF ('we have no reason to fear death' – itself rebutted by the DD-Arg), is uniformly *universal* ('we', 'all of us', 'all the people'). Notice that the denial of RF takes *existential* reading ('*some* people have *a* reason to fear death'). The DD-Arg in its *existential* reading, therefore, would formally suffice to rebut the NSH-Arg and the SYM-Arg. But notice that the SQU-Arg in its original form ('Look what *you* do with your life') takes a *singular* reading. We must hence address also its restricted form (where 'you' is restricted to 'people who waste their lives').

We first turn to the generalized form of the SQU-Arg. If all the first premises and inferences initially take the value 1, and if the scope of quantification in LV or LV+ ('our life is a valuable thing [for us]) is included in or meets the scope of SQ – so that the SQU-Arg *defeats* the DD-Arg – then also the final conclusion takes the value 1. This is so because the SQU-Arg blocks the DD-Arg, itself the only available con-argument to rebut the positive part of the argumentation. The SQU-Arg thus defeats the DD-Arg and defends the positive part of the structure.

Even if some people do *not* squander their lives, what may defeat the DD-Arg is the axiological conflict between PB ('pain is the only bad thing [for us]') and VT ('being deprived of something valuable [for us] is most certainly a bad thing [for us]'). (This conflict is covert and unmarked in Fig. 1). To appreciate the conflict, assume that pain *is* the only truly bad thing and that a bad thing is *not* valuable. If so, then being deprived of something valuable can merely be an "untruly" bad thing (i.e., not a bad thing at all), or a bad thing of a lower degree. Thus, if PB refers to a *higher*-ranked axiological value than VT does, then PB *undermines* the DD-Arg in the first case but *undercuts* the DD-Arg in the second case.

We can consider the DD-Arg not only as being *undercut* by the SQU-Arg but also as being *outweighed* by the NSH-Arg and the SYM-Arg together. This assumes that the NSH-Arg and the SYM-Arg are *stronger* than the DD-

Arg. To express this, however, requires a richer set of weights (e.g., Gordon & Walton, 2006; Selinger, 2014; and partially Tokarz, 2006) than is available in our evaluation model. This consideration also applies to aggregating the strength of convergent arguments, because the value of their sum should exceed the values of their components, i.e., the NSH-Arg and the SYM-Arg *together* should be stronger than the NSH-Arg or the SYM-Arg *separately*. But this mode of evaluation requires a separate summation of pro- and con-arguments, which makes it subject to the *double counting fallacy*. Therefore, Walton and Gordon (2015, 532) replace a summation with proof standards. Other authors, by contrast, advocate summation as a proper mode of evaluation for convergent and conductive reasoning (e.g., Govier, 1985; Hitchcock, 1983; Yanal, 1991; Selinger, 2014).

At any rate, given our exemplary evaluation model, the SYM-Arg can at most "insure" against the risk of the NSH-Arg *alone* being defeated. This, however, may not be the best insurance. For if only future events can be feared, then the analogy to the time *before* birth (itself the warrant of the SYM-Arg) becomes irrelevant, thus *rebutting* the argument that supports the premise TR ('the time before we were born is relevantly similar to the time after we die') and hence *undermining* the final inference in the SYM-Arg.

We did so far assume that ED ('we do not exist once we are dead') is analytic. This ignores the metaphysical question of whether anything else occurs if mortals die in a biological sense. Apart from the remark about the axiological conflict between PB and VT, moreover, we assumed the acceptability of all first premises. Yet, if all first premises take *universal* quantification, counterexamples to the SQU-Arg's supporting premises (CC, DA, and JA) are easy to find. It suffices to recognize that, after all, some mortals *can* sometimes stop drifting aimlessly in utter bewilderment of mind (not DA), *can* forget about the countless cares on every side (not CC), and *can* appreciate the simple joys that life offers (not JA), even give joy to others. Indeed, the *generalized* form of the SQU-Arg is readily denied by recognizing that mortals regularly do create valuable things (e.g., works of art, scientific theories, architecture). Notice, too, that the SQU-Arg's premise ST ('we sleep the greater time of our lives') would be acceptable only if 'greater' read 'great'. But even this version of the premise is undercut by the natural, biological necessity that sleep is. Add to this that sleep can be also a source of life's joys, such as pleasant dreams. Moreover, people normally help others prevent, survive, or bear 'the only bad thing' (PB) that pain allegedly is. This indeed contributes to making their lives worthwhile for others.

As similar examples undercut or directly rebut that 'we squander our lives' (SQ), the generalized form of the SQU-Arg (universal quantification) appears to be a *weak* argument. In our model of evaluation, SQ would take the value 0 or $\frac{1}{2}$, hence obtaining 0 as the value of the conclusion (RF), because the NSH-Arg and the SYM-Arg are no longer defended against the rebutting DD-Arg. Conversely, the SQU-Arg is strong if the scope of quantification is restricted to a *particular* opponent who does waste their life at *specific times*. But this demonstrates not the SQU-Arg's universal validity but merely its *occasional* dialectical validity.

So restricted, the SQU-Arg is – due to various psychological defense mechanisms – presumably ineffective vis-à-vis opponents who are yet unconvinced that their life is worthless. We investigate the SQU-Arg's persuasiveness empirically in Sect. 4.5. We will particularly inquire whether empirical data are consistent with analyzing the SQU-ARG as a dialectical response to the DD-Arg, rather than as an independent line of support for the conclusion (RF) 'death is not a bad thing to be feared'. To motivate this question, let us first turn to the Bayesian and the abstract computational approaches.

4.3. The Bayesian approach

4.3.1. Basics[2]

Interpreted subjectively, probabilities (P) do not represent objective chances of singular or repeatable events, but rather *degrees of belief* (aka credences) or *graded commitments* in reasons and claims that are mapped onto the interval $0 \leq P \leq 1$ (Korb, 2004; Hahn & Oaksford, 2007; Hahn & Hornikx, 2016; Oaksford & Hahn, 2004). The Bayesian approach applies to the reason-claim-complexes that natural language arguments express given the following abbreviations:

C: claim, conclusion, or standpoint

R: reason, or the set of conjoined premises $\{R_1 \,\&\, R_2 \,\&\, \ldots \,\&\, R_n\}$

P: probability (a measure of credence, subjective belief, or commitment)

P(R), P(C): marginal or prior probability of a reason, a claim

P(C|R): the conditional or posterior probability of a claim given a reason

P(R|C): the probability of a reason given a claim

t: an arbitrary threshold value

$\sim$: negation

Bayes' Theorem (BT) defines the posterior probability of a claim C given a reason R as the joint probability of the reason and the claim, $P(\mathrm{R\&C}) = [P(\mathrm{R}) \times P(\mathrm{C})]$, over the reason's prior probability, $P(\mathrm{R})$ (Bayes, 1763):

$$\text{(BT)} \qquad P(\mathrm{C}|\mathrm{R}) = P(\mathrm{R}|\mathrm{C}) \times P(\mathrm{C})/P(\mathrm{R})$$

To obtain $P(\mathrm{C}|\mathrm{R})$, interpreted as the *posterior* probability of C given R, requires multiplying the claim's *prior* probability, $P(\mathrm{C})$, with $P(\mathrm{R}|\mathrm{C})/P(\mathrm{R})$. Call this latter term the *impact* of the reason on the claim, i. This term decomposes as:

$$\begin{aligned}\text{(Impact term)} \qquad i &= P(\mathrm{R}|\mathrm{C})/P(\mathrm{R}) \\ &= P(\mathrm{R}|\mathrm{C})/[(P(\mathrm{R}|\mathrm{C}) \times P(\mathrm{C}) + P(\mathrm{R}|{\sim}\mathrm{C}) \times P({\sim}\mathrm{C})]\end{aligned}$$

Since i is the ratio of how probable the reason is *given* the claim to how probable the reason is *marginally* (irrespective of the claim), i expresses a conditional expectation of the reason if the claim holds, as against a prior expectation on the reason regardless. (We return to the impact term below.)

The use of BT presupposes a *Pascalian* approach to probability, where changes in credence in response to reasons for or against C entail changes regarding its negation ∼C, defined as $P({\sim}\mathrm{C}) = 1 - P(\mathrm{C})$. This definition also holds for the priors, which express commitments to the truth or falsity of C given background information, as well as for conditional probabilities, $P(\mathrm{R}|\mathrm{C}) = 1 - P({\sim}\mathrm{R}|\mathrm{C})$ and $P(\mathrm{R}|{\sim}\mathrm{C}) = 1 - P({\sim}\mathrm{R}|{\sim}\mathrm{C})$. The terms $P(\mathrm{R}|\mathrm{C})$ and $P(\mathrm{R}|{\sim}\mathrm{C})$ both denote a *likelihood*, i.e., a probability multiplied by a positive constant (Edwards et al., 1963), expressing prior commitments regarding the probative value of a reason for a claim. (For the contrasting *Baconian* approach to probability, see Cohen, 1980; Spohn, 2012; Zenker, 2015, Sect. 5.3).

4.3.2. Application to natural language argument

The term $P(\mathrm{R}|\mathrm{C})$ expresses the reason's *sensitivity* to the claim. For instance, when evaluating the reliability of an empirical test – where 'hypothesis (H)' replaces 'claim' and 'evidence (E)' replaces 'reason' – $P(\mathrm{R}|\mathrm{C})$ reports the *true positive rate* (the ratio of *correct* positive test-results and all test-results), and $P(\mathrm{R}|{\sim}\mathrm{C})$ denotes the complement of the *specificity* of the reason to the claim, i.e., the *false positive rate* (the ratio of *incorrect* positive test-results and all test-results). If long-run frequencies of token-events

leave both terms interpretable, then sensitivity and specificity together express the degree to which R and C correlate. In the extremal cases, if C logically entails R, then $P(\mathrm{R}|\mathrm{C}) = 1$, and, if statistical independence holds, then $P(\mathrm{R}|\mathrm{C}) = P(\mathrm{R}|{\sim}\mathrm{C}) = P(\mathrm{R})$.

In contexts where reasons support or undermine claims *irrespective* of frequency considerations, both likelihoods remain meaningful on the following interpretation: reason R is *sensitive* to claim C to the extent that R supports C more than R supports any other claim C* (entailing ∼C), i.e., $P(\mathrm{C}|\mathrm{R}) > P({\sim}\mathrm{C}|\mathrm{R})$; and R is *specific* to C to the extent that R rather than any other reason R* (entailing ∼R) supports C, i.e., $P(\mathrm{C}|{\sim}\mathrm{R}) < P({\sim}\mathrm{C}|{\sim}\mathrm{R})$. When drawing sensitivity and specificity together, the extent to which R supports C thus depends on the extent to which the *C*-supporting-reason R *fails* to support ∼C, as well as on the extent to which argumentative support for C *cannot* be generated by any reason besides R (Godden & Zenker, 2018). Argumentative support therefore is strongest if R is an *exclusive and decisive* reason-for-C, and weakest if R is a *common and indecisive* reason-for-C. An exclusive and decisive reason can be likened to the application of an *ideal* test for which no alternative test is available. For instance, assume (unrealistically) that the test is perfectly sensitive, $P(\mathrm{R}|\mathrm{C}) = 1$, and also perfectly specific, $P(\mathrm{R}|{\sim}\mathrm{C}) = 0$. So, if $P(\mathrm{R}) = 1$, and if $P(\mathrm{C}) = 0.5$ (marking the undecided case), then evidence (reasons) can support a hypothesis (claim) *decisively* because, given $i = P(\mathrm{R}|\mathrm{C})/P(\mathrm{R})$, we find that $i = 2.0$, whence $P(\mathrm{C}|\mathrm{R}) = i \times P(\mathrm{C}) = 2 \times 0.5 = 1$.

Given prior probability values for R and C, values for sensitivity and selectivity thus suffice to model *argument strength* as the comparative support that R provides to C, for instance as $P(\mathrm{C}|\mathrm{R}) - P(\mathrm{C})$. Other measures remain possible (Pfeifer, 2013). The final value of $P(\mathrm{C}|\mathrm{R})$ can also be interpreted as the *force* of the argument (Oaksford & Hahn, 2004). Moreover, a context-sensitive threshold t can precisify commitments to the 'is a necessary/insufficient/sufficient/supererogatory reason for'-relation (Spohn, 2012), respectively to precisify contextual constraints on relevance, sufficiency, and acceptability (Johnson & Blair, 1977), such as apply in statistics for *evidential strength* or in criminal law for *evidential value* (Godden & Zenker, 2018).

Crucially, unless frequentist information grounds these probability values, they depend on what arguers believe or commit to, and so depend on what arguers recognize as (strong) reasons. A constraint other than assigning and updating probability values *coherently*, as BT dictates, is not a part of the Bayesian approach.

4.3.3 What a Bayesian analysis presupposes

Rather than giving a Bayesian analysis of the SQU-Arg, notice that any such analysis presupposes non-zero (point or interval) numerical values to represent prior degrees of commitment associated with premises and conclusion. If non-zero values for likelihoods are in place, too, then the support that any number of premises lend to a conclusion can be calculated by applying BT *iteratively*, i.e., by conditionalizing the conclusion on the premises in any order.

For instance, one might conditionalize the conclusion, S, first on the NSH-Arg, then on the SYM-Arg, and finally on the SQU-ARG, to derive P(S|NSH, SYM, SQU). If the three arguments increase support for S *incrementally* as per BT, each argument provides an *independent* line of support (Fig. 2, left pane). Whereas if support is lower than what BT states for P(S|NSH, SYM, SQU), then at least one argument is not independent (Fig. 2, right pane), precisely as the dialectical analysis had suggested for the SQU-Arg. In this case, the support for S is given as P(S|NSH, SYM, DD), where P(DD) = P(DD|SQU). Because the case that obtains depends on how *arguers themselves* assign numerical values, it translates into the empirical question addressed in Sect. 4.5.

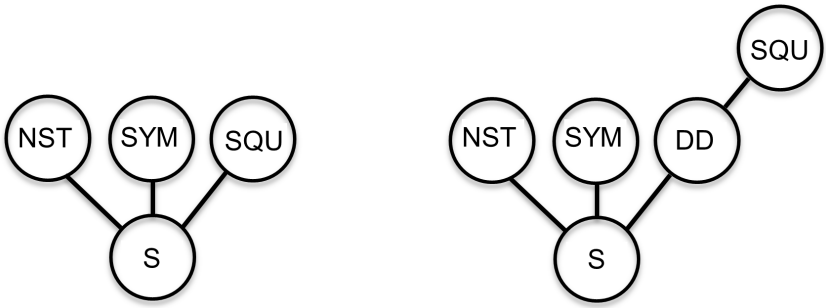

Figure 2. The SQU-Arg as an independent line of support for the conclusion S (left), or as a dialectical response to the DD-Arg (right)

Beyond the coherence constraint, a Bayesian approach could dictate argument strength only if specific numerical values were "right" – which is a hard-to-motivate assumption – and if the analysis as *independent* support lines were infallible – which is generally not the case. A Bayesian approach primarily provides a modeling tool that captures argument strength as *coherent* arguers would apply it. The application of this tool thus depends on the non-trivial task of getting the argument structure right (see Bex & Renooij, 2016; Wieten, Bex, Prakken & Renooij, 2019).

The same holds for the computational approach to argument strength.

4.4. The computational approach

4.4.1. Conflicts, semantics, and three notions of argument strength

Among the range of computational approaches to argumentation, a computational analog can be provided for each approach to argument strength discussed so far. In this section, we nevertheless focus on the classical abstract argumentation approach by Dung (1995; see Pollock, 1995). A Dung-style approach abstracts from the statements expressing an argument and how they work together to strengthen or weaken a position. This approach subsumes premise-conclusion structures into distinct, atomic, abstract arguments. Although an argument's internal premises-conclusion structure is retained, the focus shifts from the support *within* an argument to the conflict *between* arguments. In this, analysts support the computational machinery by making leaps of intuition and teasing out meaning from unexpressed or under-formed text, which machines cannot yet do by themselves.

Studying undercutting and rebutting attacks (Walton, 2009) between abstract arguments allows computing an argument's acceptability-status relative to a pattern of conflicts in a directed graph known as an *argumentation framework*. For instance, if argument A1 is attacked by argument A2, and A2 is itself attacked by A3, then A3 defends A1. This also applies to sets of arguments. An argument is acceptable *with respect to a set of arguments* if it is defended by a member of that set. An entire set of arguments is *conflict-free* if there are no attacks between its members. An *admissible set* can thus be defined as a set of conflict-free arguments the members of which are reflexively acceptable (i.e., acceptable to all other set members).

The properties of *conflict-freeness* and *acceptability*, as well as the derived property of *admissibility*, are central in identifying specific subsets of an argumentation framework. This is often referred to as *applying a semantics*. A semantics provides ways of understanding and interpreting an argumentation framework. A semantics thus is an evaluation determining which consistent groups of arguments – called *extensions* – can be accepted. For example, the *stable* semantics seeks to identify the unambiguous extensions for a given framework but, in some cases, will yield no result. Whereas the *semi-stable* semantics seeks to account for the largest possible set of arguments in the framework.

Because there can be multiple extensions, the extension one would best adopt as a position should be sensitive to an argument's strength. Although the computational approach rarely makes explicit reference to argument

strength, one can identify three related concepts of argument strength: (i) based on the relation between sets of arguments calculated by a given semantics, (ii) based on the relationship between different semantics, and (iii) based on using preferences to order the members of an argumentation framework.

The *first* concept rests upon the output of a semantics by applying the so-called "gunfighter" analogy: arguments (gunfighters) that are acceptable (survive) when the evaluation process (gunfight) terminates are strong because they are undefeated. Given a framework and a specified semantics, an *acceptable* argument thus is objectively strong.

The *second* concept involves comparing the set of arguments identified by some semantics as stronger or weaker than those identified by another semantics. Given many semantics, as well as an equivalent (slightly more expressive) approach called *labeling* (Verheij, 1996), these approaches all fit within a hierarchical organization (Baroni et al., 2018). A stable semantics generally instantiates a semi-stable semantics, which in turn instantiates a *preferred* semantics, with the conflict-free set as the base condition. Differences in argument strength may thus be derived from the differences in requirements for set membership under a semantics. And members of an extension under a complete semantics, for instance, may be claimed to be *stronger* than members of a merely admissible set.

But what persuades one person might not persuade another person and may even be counterproductive. An individual "gunfighter"-strong argument therefore need not be persuasive in general. As neither the first nor the second concept of argument strength can identify the strongest individual argument within the extension, the *third* concept recognizes a subjective sense of argument strength resting on preference orderings over the arguments within a framework (Amgoud & Cayrol, 2002). This reflects that people are convinced for different (debatable) reasons. Although a general method to construct an accurate preference ordering for an audience is unavailable, techniques like those used in the empirical approach would be a good starting point. Regardless, if a set of arguments is acceptable, then the *strongest* argument could be identified as the *most preferred* set member.

4.4.2. A computational analysis of the squandering argument

On the most basic interpretation, the computational approach uses extant approaches to natural language argument as a foundation to construct a system that automatically processes arguments. The first of two levels of relationship between the computational and extant approaches is

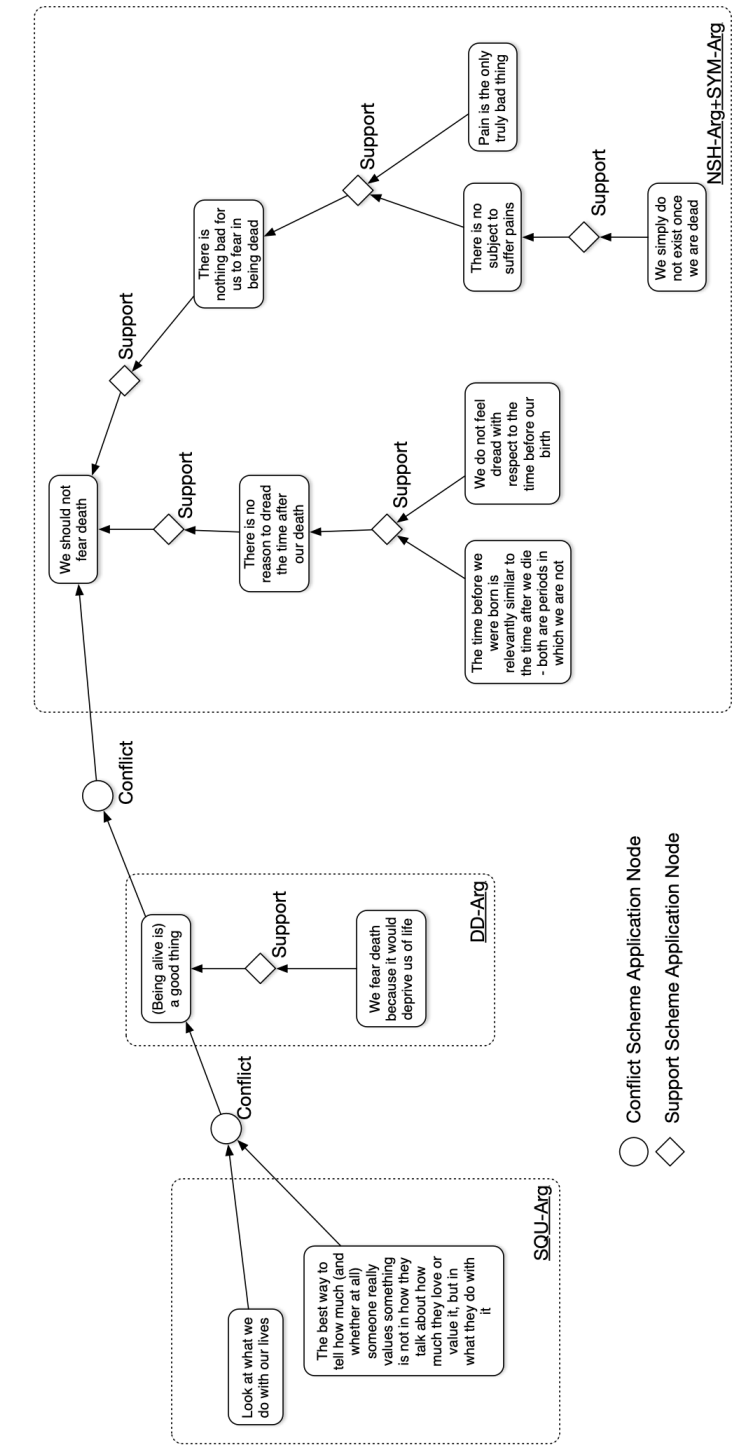

Figure 3. Abstract argument network

constitutive, concerning the supply of structured data to construct a framework of arguments and dialectical relations. This relates to argument strength via Dung's (1995) analysis of conflicts. The second level is *evaluative*, concerning mechanisms to arrange such arguments (e.g., via a preference ordering).

Regarding the text of Aikin & Talisse (2019), the fine structure of the arguments must first be identified, including premise-conclusion-relations, the schemes instantiated, and more complex structures arising from combining simple ones (e.g., as serial, con-/divergent, or complex forms). Subsequently, a dialectical analysis of conflicts amongst the arguments can be performed. An analysis that stays close to Aikin & Talisse's (2019) text yields the diagram in Fig. 3. Of course, the use of enthymematic argument, the omission of argument parts, or entire dialectical paths, can lead to a discrepancy between what a source text states and what an analyst makes of it.

The abstract framework in Fig. 4 is derived from the argument analysis in Fig. 3. The support structures of the four arguments are abstracted into three abstract argument *nodes*, the NSH-Arg+SYM-Arg, the DD-Arg, and the SQU-Arg. Notice that both the the NSH-Arg and the SYM-Arg are subsumed into a single abstract node, as both represent independent lines of support for the same conclusion S ('we should not fear death'). Attacks between these arguments are depicted by arrows such that the SQU-Arg attacks the DD-Arg, and the DD-Arg attacks the NSH-Arg+SYM-Arg.

Figure 4. Abstract argument framework based on the analysis in Fig. 3

Initially, the NSH-Arg+SYM-Arg is defeated by the attack from the DD-Arg, which is in turn defeated by the SQU-Arg. This causes the NSH-Arg+SYM-Arg to be reinstated in the absence of further undefeated attacks. On a Dung-style approach, the prevailing argument thus is the NSH-Arg+SYM-Arg ('we should not fear death'). But as the NSH-Arg+SYM-Arg and SQU-Arg both survive (absent further attacks), they are of equitable strength, i.e., there is no basis to assess the NSH-Arg+SYM-Arg as stronger than the SQU-Arg, nor can it be considered stronger than the DD-Arg which was defeated.

4.5. The empirical approach

4.5.1. Assessing argument strength empirically

In the branch of communication studies known as persuasion research, argument strength is empirically assessed using evaluation methods based on a population's cognitive responses to experimentally manipulated message contents. Indirect assessment methods such as the open-ended method of *thought listing* (developed in Petty & Cacioppo's (1986) elaboration likelihood model) allow researchers to compare whether independent empirical results converge. Thought listing is rightly criticized as an inadequate indicator of argument strength, because the valence of a population's thoughts (aka thought-favorability) is interpreted as a *single* dimension of argument strength (Munch & Swasy, 1988, Stephenson & Palmgreen, 2001; Zhao & Cappella, 2016). A second limitation is the potential lack of capacity or felt need to provide *precise* thoughts. This particularly affects responses from students, who are typically recruited as experimental participants (Henrich, Heine, & Norenzayan, 2010).

Argument strength indicators beyond a single dimension of thought-favorability are provided by closed-ended methods that rely on multi-item scales of perceived argument strength. Based on work by Munch & Swasy (1988) and Lavine & Snyder (1996) (see Carpenter, 2015; Darke & Chaiken, 2005), the present investigation adopted an argument strength scale by Zhao et al.'s (2011), developed in two empirical studies with adolescent and adult participants who assessed the strength of arguments used in public service announcements that discourage drug use and smoking. This scale comprises *nine* indicators: believability, novelty, convincingness, importance, confidence, friend, thoughts, agreement, and reason (see the next subsection).

Allowing for plausible error correlations, Zhao et al. (2011) established the scale's descriptive adequacy vis-à-vis participant's cognitive responses by removing indicators that displayed the lowest loadings on the latent factor of argument strength: *convincingness* and *novelty* in the case of drug use, and *novelty* in the case of smoking. Since both of Zhao et al.'s (2011) studies failed to show that *novelty* is a relevant indicator, our evaluation of the perceived strength of the NSH-Arg, SYM-Arg, and SQU-Arg relies on the remaining eight indicators. This scale served to calculate the correlations between an argument strength index and a thought favorability index.

Crucially, participants were presented with all three arguments as *independent* lines of support for the conclusion (RF) 'death is not a bad thing to be feared', although using the formulation 'there is no reason to fear

death'. This presentation is *inconsistent* with analyzing the SQU-Arg as a dialectical response to the DD-Arg (see Fig. 2). Markedly lower argument strength-ratings/correlations for the SQU-Arg compared to the NSH-Arg and SYM-Arg can therefore be interpreted as indirect evidence *consistent* with the dialectical analysis above.

4.5.2. Method

Participants were 52 native speakers of Polish, commanding high or very high English proficiency (C1–C2; *Common European Framework of Reference for Languages*), who studied English Philology or English Studies: Literature and Culture for a BA or MA degree at the Faculty of English at Adam Mickiewicz University, Poznań, Poland (age M = 21.2, SD = 1.4; age range 19–26 years; 10% male, 90% female). Volunteering to exercise their critical/evaluative skills in the English language, participants were presented with a cover story purporting that research addressed how fear of different objects influences language processing. Participants were asked to indicate fear-related attitudes to twelve 7-point Likert scale questionnaire items (1 = not afraid at all, 7 = very much afraid). Eleven items were fillers; the sixth item asked for their attitude to death.

Participants indicated their agreement with three sets of statements corresponding to the premises and the deductive inferential link between the premises and the conclusion of the NSH-Arg, the SYM-Arg, and the SQU-Arg. The statements for the *NSH-Arg* read: (A) Pain is the only truly bad thing; (B) We do not feel pain once we are dead, and (C) IF pain is the only truly bad thing, AND we do not feel pain once we are dead, THEN there is nothing bad for us to fear in being dead. The statements for the *SYM-Arg* read: (A) We do not feel dread with respect to the time before our birth; (B) The time before we were born is relevantly similar to the time after we die, and (C) IF we do not feel dread with respect to the time before our birth, AND the time before we were born is relevantly similar to the time after we die, THEN there is no reason to dread the time after our death. And the statements for the *SQU-Arg* read: (A) A sufficient reason to fear death entails not wasting one's life; (B) People waste their lives; and (C) IF a sufficient reason to fear death entails not wasting life, AND people waste life, THEN people have no sufficient reason to fear death.

Participants expressed their agreement with each set of statements as a rating for *believability*, *convincingness*, *importance*, *confidence*, *friend*, *thoughts*, and *agreement* using a five-point scale (1 = strongly disagree, 5 = strongly agree) and also expressed a rating for *reasons* (1 = very weak, 5 = very strong). The procedure asked to assess the extent to which

(1) statements A and B are *believable* reasons for not being afraid of death; (2) statements A and B are *convincing* reasons for not being afraid of death; (3) statements A and B are reasons for not being afraid of death that are *important* to me; (4) statements A and B *helped me feel confident* about the reasons for not being afraid of death; (5) statements A and B would *help my friends* not be afraid of death; (6) statements A and B put *thoughts* in my mind *about not being afraid* of death; (7) statements A and B put *thoughts* in my mind *about being afraid* of death; (8) overall, how much do you *agree or disagree* with statements A and B?; (9) are the *reasons* that statements A and B give for not being afraid of death *strong or weak*?

4.5.3. Results

Based on assessments of statements A and B, the *thought favorability index* (range: –4 to +4) was arrived at by subtracting the score for unfavorable thoughts (afraid of death) from the score for favorable thoughts (not afraid of death). This index was converted to a five-point scale (dividing the final scores by 2 and adding the constant 3) to agree with other scale items. We also measured each argument's comprehensibility, complexity, and familiarity on a 5-point scale (1 = *strongly disagree*, 5 = *strongly agree*) (Table 1).

To evaluate participants' attitudes to the reasons each argument provides for the claim 'there is no reason to fear death', participants were asked to note their thoughts on these reasons (Cacioppo & Petty, 1981; see Zhao et al., 2011). The expectation is that favorable, respectively unfavorable, thoughts arise as responses to a message perceived as strong, respectively as weak. Two researchers independently coded the responses as *favorable*, *neutral*, or *unfavorable*. (Disagreements were resolved by discussion). This generated a thought favorability index by subtracting, for each participant, the sum of unfavorable thoughts from that of favorable thoughts, followed by averaging (Table 1).

Table 1

Thought favorability, comprehensibility, complexity, and familiarity

	Thought favorability	comprehensibility	familiarity	complexity
NSH-Arg	– 1.19 (.25)	3.69 (1.18)	3.16 (1.28)	2.06 (1.02)
SYM-Arg	– 0.66 (.25)	3.21 (1.16)	2.38 (1.23)	2.96 (1.18)
SQU-Arg	– 0.68 (.22)	3.02 (1.19)	2.83 (1.27)	3.04 (1.22)

Note: values report means on a five-point scale, with standard deviations in brackets.

For the NSH-Arg, the full eight-item scale (*thoughts* plus *believability*, *convincingness*, *importance*, *confidence*, *friend*, *agreement*, *reasons*) displayed good internal consistency ($\alpha = .84$). The α-value slightly increased after removing the *thoughts* ($\alpha = .85$) or the *friend* ($\alpha = .86$) indicator – thus identifying both as the weakest indicators of perceived argument strength – whereas the full scale already displayed an excellent internal consistency ($\alpha = .94$) for the SYM-Arg. The α-value again increased after removing the *thoughts* or *friend* indicator (both $\alpha = .95$). The SQU-Arg, by contrast, only reached $\alpha = .86$, and increased but slightly when removing the *thoughts* or *friend* indicator (both $\alpha = .87$).

To determine the indicators that best represent perceived argument strength, a confirmatory factor analysis relied on a comparative fit index (CFI), the root mean square error of approximation (RMSEA) and a chi-square (χ^2) test. A CFI $\geq .95$ and an RMSEA $\leq .06$ are considered acceptable (Zhao et al., 2011, Brown & Moore, 2012). As above, the original scale failed to fit well for the NSH-Arg (χ^2 (20, N = 52) = 32.01, $p = .04$, CFI = .92, RMSEA = .10, 90% CI [.01, .17]). Standardized regression weights were low for the indicators *friend* (.30) and *thoughts* (.38), relatively high for *reasons* (.58), and high for all other indicators (90% CI [.71, .98]). Perfect fit resulted after removing *friend* (χ^2 (14, N = 52) = 20.34, $p = .12$, CFI = .95, RMSEA = .09, 90% CI [.00, .17]) and *thoughts* (χ^2 (9, N = 52) = 5.09, $p = .82$, CFI = 1.0, RMSEA = .00, 90% CI [.00, .09]). For the SYM-Arg, the original scale (χ^2 (20, N = 52) = 25.46, $p = .18$, CFI = .98; RMSEA = .07, 90% CI [.00, .14]) again fitted well after removing *thoughts* (.43) and *friend* (.68) (χ^2 (9, N = 52) = 11.15, $p = .26$, CFI = .99, RMSEA = .06, 90% CI [.00, .17]). For the SQU-Arg, by contrast (χ^2 (20, N = 52) = 59.70, $p < .001$, CFI = 0.78, RMSEA = .20, 90% CI [.04, .29]), removing *thoughts* (.34) and *friends* (.40) did improve the CFI-value. Yet this decreased the RMSEA-value but insufficiently, and so failed to fit data (χ^2 (5, N = 52) = 12.16, $p = .03$, CFI = .93, RMSEA = .16, 90% CI [.04, .29]) (Table 2).

Table 2

Perceived argument strength index

	Perceived argument strength	indicators removed	CFI	RMSEA	RMSEA 90% CI
NSH-Arg	2.71 (.87)	friends, thoughts	1.0	.00	[.00, .09]
SYM-Arg	2.48 (1.22)	friends, thoughts	.99	.06	[.00, .17]
SQU-Arg	n/a	friends, thoughts	.93	.16	[.04, .29]

Note: values report means on a five-point scale, with standard deviations in brackets; 'n/a' = not applicable as RMSEA too high and CFI too low; 'CI' = confidence interval

For all three arguments, an analysis of variance for repeated measures did not result in significant differences in thought-favorability ($F(2, 92) = 2.13$, $p = .12$) (Table 1). The same holds for pairwise comparisons: NSH-Arg/SYM-Arg ($p = .32$); NSH-Arg/SQU-Arg ($p = .18$). Moreover, internal correlations between perceived strength and thought-favorability were *moderate* for the NSH-Arg ($r = .47$, $p = .001$), *strong* for the SYM-Arg ($r = .66$, $p = .001$), but *low* for the SQU-Arg ($r = .33$, $p = .020$).

4.5.4. Discussion

Zhao et al.'s (2011) original eight-item scale, though well-grounded in the literature, failed to fit the data for all three arguments. Removing the *thoughts* and the *friend* indicator resulted in an acceptable descriptive model of cognitive responses to the NSH-Arg and the SYM-Arg. But this model did not fit the data for the SQU-Arg. The *thoughts* indicator even decreased the original scale's descriptive adequacy for the NSH-Arg and the SYM-Arg. The (un)favorable thought-contents (elicited through thought-listing) nevertheless correlated *strongly* with perceived strength ratings for the SYM-Arg, *moderately* for the NSH-Arg, and *weakly* for the SQU-Arg. This implies that these open- and closed-ended evaluation methods converge for the NSH-Arg and the SYM-Arg, but not for the SQU-Arg.

We submit this as *indirect* evidence consistent with having presented all three arguments to participants as *independent* lines of support for the claim 'there is no reason to fear death', a presentation that *failed* to preserve the dialectical character of the SQU-Arg as an argumentative response to the DD-Arg. Conversely, if the contrasting dialectical analysis of the SQU-Arg as a dialectical response is accepted, then this result is expectable. A future study seeking *direct* evidence for the dialectical analysis should therefore present the arguments according to this analysis.

5. Conclusion

On the computational approach, the NSH-Arg, the SYM-Arg, and the SQU-Arg prevail simply because the SQU-Arg *defends* the DD-ARG. But without further information, nothing else can be said about the arguments' comparative strengths. Similarly, without *numerical values* to calculate the support for a conclusion, nothing substantial about argument strength comes forth on a structural or a Bayesian approach. And what Bayes' theorem or the evaluation mechanism of the structural approach dictates about argument strength depends on analysts getting the argument structure right first.

Although the required numerical values are provided by the empirical approach, it too requires a prior analysis of the argument structure. Differences in argument structure were discernible in the judgments of perceived argument strength measured with Zhao et al.'s (2011) multi-item scale. The scale allowed for an evaluation of the strength of the NSH-Arg and the SYM-Arg but failed to fit the data for the SQU-Arg. This was presumably because participants were presented with the SQU-Arg as an *independent* line of support, ignoring its status as a dialectical response. A follow-up study should present the SQU-Arg according to our dialectical analysis.

Given assumptions, the dialectical analysis allowed a clear verdict on the strength of the SQU-Arg *at a specific point* in the dialogue. Once PRO raised the SQU-Arg, and assuming no objection by RES, a winning strategy was not apparent for either discussant. RES's one available move was to *concede* PRO's standpoint by retracting not-S ('we should fear death') because RES retains an *unfulfilled* obligation to defend not-S. At this point, the NSH-Arg stood unretracted and thus prevailed. But lack of information kept from evaluating how RES might defend not-S, or raise critical doubt against the NSH-Arg or the SYM-Arg.

The SQU-Arg's *generalized* form was clearly undermined (e.g., by valuable things like works of art, scientific theories, or architecture) and potential *defeaters*, whose impact on argument strength can be investigated empirically, were also available (e.g., sleep as a biological necessity or as a source of life's joys). The dialectic approach can represent these defeaters by extending the move space, whereas the Bayesian, structural, or computational approaches require a revision of the argument structure. Even if the NSH-Arg prevails, PRO was committed to the premises and the reasoning of the LGT-Arg, and thus to (11): 'in dying we get less of a good thing'. But the dialogue did not indicate how PRO might undercut the inference from (11) to not-S. Specifying this undercutter thus remains an unfulfilled burden for PRO.

Finally, a dialectical analysis includes *premise assertability* as a standard, thus addressing the pragmatic dimension of arguing. An acceptable premise can therefore be deemed "faulty" if a proponent cannot meet their obligation to defend it *qua* lacking a *right* to assert a defense, obliging them to retract an otherwise supportable claim. This evaluation can only be reached by integrating the pragmatic dimension. When evaluated dialectically, the SQU-Arg therefore is a non-fallacious *ad hominem* argument, although it is not necessarily stronger than possible arguments attacking it.

Acknowledgments

This is a substantially revised version of the authors' contributions to the *Symposium on Argument Strength* at the 3rd European Conference on Argumentation (ECA), held 24–27 June 2019 in Groningen, The Netherlands (see Zenker et al., 2020). All authors contributed to drafting the manuscript, which FZ edited. Additional language editing was provided by Simon Wells. All authors approved the final submitted version. F. Z. and K. D-K. acknowledge support from the European Network for Public Policy Argumentation (COST Action CA17132).

Data Availability Statement

The materials and data from this study are available on request from the corresponding author, [K. D-K].

NOTES

[1] Since supplementing the premises of a defeasible inference with the negatives of its undercutters strengthens this inference, systems equipped with richer sets of weights can use these to express the relevance of undercutters (Selinger, 2019). But the exemplary system outlined here must define relevant undercutters by explicitly stating all exceptions to defeasible rules, e.g., in the meta-language (*cf.* Modgil & Prakken, 2014, p. 35).

[2] Section adapted from Godden & Zenker (2018), who explicate formal steps omitted here and provide further references (see Zenker, 2013).

REFERENCES

Aikin, S. F., & Talisse, R. B. (2019). Epicureans on squandering life. *Three Quarks Daily*. URL=https://www.3quarksdaily.com/3quarksdaily/2019/02/epicureans-on-squandering-life.html (accessed 10 FEB 2022).

Aikin, S. F., & Talisse, R. B. (2022). Epicureans on Death and Lucretius' Squandering Argument. *Southwest Philosophy Review*, 38(1), 41–49.

Amgoud, L., & Cayrol, C. (2002). A reasoning model based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence*, 34 (1–3), 197–215.

Barth E., & Krabbe, E. C. W. (1982). *From axiom to dialogue*. Berlin: de Gruyter.

Baroni, P., Caminada, M., & Giacomin, M. (2018). Abstract argumentation frameworks and their semantics. *Handbook of Formal Argumentation* (chapter 4, pp. 159–236.) London: College Publications.

Bayes, T. (1763/1958). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53, 370–418. (Reprinted in *Biometrika*, 45, 296–315.)

Bex, F. J. & Renooij, S. (2016). From Arguments to Constraints on a Bayesian Network. Computational Models of Argument. In P. Baroni, T. F. Gordon, T. Scheer and M. Stede (eds.) (2019), *Proceedings of COMMA 2016, Frontiers in Artificial Intelligence and Applications, volume 287* (pp. 96–106). IOS Press: Amsterdam.

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 361–379). New York, N.Y.: The Guilford Press.

Darke, P. R., & Chaiken, S. (2005). The pursuit of self-interest: Self-interest bias in attitude judgment and persuasion. *Journal of Personality and Social Psychology*, 89, 864–883.

Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in non-monotonic reasoning and logic programming and n-person games. *Artificial Intelligence*, 77, 321–357.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review*, 70(3), 193–242. http://dx.doi.org/10.1037/h0038674

Finocchiaro, M. (2007). Arguments, meta-arguments, and metadialogues: A reconstruction of Krabbe, Govier, and Woods. *Argumentation*, 21, 253–268.

Finocchiaro, M. (2013). *Meta-argumentation: An approach to logic and argumentation theory*. London: College Publications.

Godden, D., & Zenker, F. (2018). A probabilistic analysis of argument cogency. *Synthese*, 195, 1715–1740.

Gordon, T. F., & Walton, D. (2006). The Carneades argumentation framework. In P. E. Dunne & T. J. M. Bench-Capon (Eds.), *Computational Models of Argument: Proceedings of COMMA 2006* (pp. 195–207). Amsterdam: IOS Press.

Govier, T. (1985). *A practical study of argument*, 1st through 7th editions. Belmont, CA: Wadsworth.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: a bayesian approach to reasoning fallacies. *Psychological Review*, 114, 704–732.

Hahn, U., & Hornikx, J. (2016). A normative framework for argument quality: Argumentation schemes with a Bayesian foundation. *Synthese*, 193, 1833–1873.

Hamblin, C. (1970). *Fallacies*. London: Methuen.

Hamblin, C. (1971). Mathematical models of dialogue. *Theoria*, 37, 130–155.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and brain sciences*, 33(2–3), 61–83.

Hitchcock, D. (1983). *Critical Thinking: A Guide to Evaluating Information*. Toronto: Methuen.

Johnson, R. H., & Blair, J. A., (1977). *Logical self-defense*. Toronto: McGraw-Hill Ryerson.

Keiff, L. (2011). Dialogue logic. In E. Zalta et al. (Eds.), *Stanford encyclopedia of philosophy*. Stanford, CA: Center for the Study of Language and Information. http://plato.stanford.edu/archives/sum2011/entries/logic-dialogical/ (accessed 5 OCT 2019)

Kononova, A., Yuan, S., & Joo, E. (2017). Reading about the flu online: How health-protective behavioral intentions are influenced by media multitasking, polychronicity, and strength of health-related arguments. *Health Communication*, 32(6), 759–767.

Korb, K. (2004). Bayesian informal logic and fallacy. *Informal Logic*, *24*, 41–70.

Krabbe, E. C. W. (1992). So what? Profiles for relevance criticism in persuasion dialogue. *Argumentation*, 6, 271–283.

Krabbe, E. C. W. (1999). Profiles of dialogue. In J. Gerbrandy, M., Marx, M. de Rijke, & Y. Venema (Eds.), *JFAK: Essays dedicated to Johan van Benthem on the occasion of his 50th birthday* (pp. 25–36). Amsterdam: Amsterdam University Press.

Krabbe, E. C. W. (2002). Profiles of dialogue as a dialectical tool. In F.H. van Eemeren (Ed.), *Advances in pragma-dialectics* (pp. 153–167). Amsterdam–Newport News, VA: Sic Sat/Vale Press.

Krabbe E. C. W. (2003). Metadialogues In: van Eemeren F.H., J.A. Blair, C. Willard C., and A.F. Snoeck Henkemans (Eds.), *Anyone who has a view* (pp. 83–90). Dordrecht: Springer.

Krabbe, E. C. W. (2013). The topical roots of formal dialectic. *Argumentation*, 27, 71–87.

Krabbe, E.C.W. (2017). The formalization of a critical discussion. *Argumentation*, 31, 101–119.

Krabbe, E. C. W., & Walton, D. (2011). Formal dialectical systems and their uses in the study of argumentation. In E. Feteris, B. Garssen, & A.F. Snoeck Henkemans (Eds.), *Keeping in touch with pragma-dialectics* (pp. 245–263). Amsterdam: John Benjamins.

Lavine, H. & Snyder, M. (1996). Cognitive Processing and the Functional Effect in Persuasion: The Mediating Role of Subjective Perceptions of Message Quality. *Journal of Experimental Social Psychology* 32(6), 580–604.

Modgil, S., & Prakken, H. (2014). The ASPIC+ framework for structured argumentation: a tutorial. *Argument and Computation*, 5(1), 31–62.

Munch, J. M., & Swasy, J. L. (1988). Rhetorical question, summarization frequency, and argument strength effects on recall. *Journal of Consumer Research*, 15, 69–76.

Oaksford, M., & Hahn, U. (2004). A Bayesian approach to the argument from ignorance. *Canadian Journal for Experimental Psychology*, 58, 75–85.

Prakken, H. (2010). An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1, 93–124.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to persuasion.* New York, NY: Springer-Verlag.

Pfeifer, N. (2013). On argument strength. In F. Zenker (Ed.), *Bayesian argumentation: The practical side of probability* (pp. 185–193). Dordrecht: Springer.

Pollock, J. L. (1987). Defeasible reasoning. *Cognitive Science*, 11, 481–518.

Pollock, J. L. (1995). *Cognitive carpentry: A blueprint for how to build a person.* Cambridge, MA: The MIT Press.

Quine, W. V. O. (1961). Two dogmas of empiricism. In *From a logical point of view* (pp. 20–46). Cambridge, MA: Harvard UP.

Quine, W. V. O. (1992). *Pursuit of truth.* Cambridge, MA: Harvard UP.

Rescher, N. (1977). *Dialectics: A controversy-oriented approach to the theory of knowledge.* Albany, NY: SUNY Press.

Selinger, M. (2014). Towards formal representation and evaluation of arguments. *Argumentation*, 28(3), 379–393.

Selinger, M. (2019). Towards defining the relation of attack. In B. Garssen, D. Godden, G. R. Mitchell, & J. H. M. Wagemans (Eds.), *Proceedings of the Ninth Conference of the International Society for the Study of Argumentation* (pp. 1064–1073). Amsterdam: Sic Sat.

Spohn, W. (2012). *The Laws of Belief.* Oxford: Oxford University Press.

Stiff, J. B. & Mongeau, P. A. (2016). *Persuasive communication.* New York: The Guilford Press.

Thomas, S. N. (1986). *Practical reasoning in natural language.* Prentice-Hall, Englewood Cliffs, New Jersey. [First edition 1973.]

Tokarz, M. (2006). *Argumentacja, perswazja, manipulacja* [*Argumentation, persuasion, manipulation*]. Gdańsk: Gdańskie Wydawnictwo Psychologiczne.

van Eemeren, F. H., & Grootendorst, R. (2004). *A systematic theory of argumentation.* Cambridge: Cambridge UP.

van Eemeren, F.H., Garssen, B., Krabbe, E.C.W., Snoeck Henkemans, A.F., Verheij, B., & Wagemans, J. (2014). Formal dialectical approaches. In van Eemeren et al. (Eds.) *Handbook of argumentation theory* (ch. 6, pp. 301–372). Dordrecht: Springer.

von Borgstede, C., Andersson, M., & Hansla, A. (2014). Value-congruent information processing: The role of issue involvement and argument strength. *Basic and Applied Social Psychology*, 36(6), 461–477.

Verheij, B. (1996). Two approaches to dialectical argumentation. Admissible sets and argumentation stages. *Proceedings of the eighth Dutch conference on artificial intelligence* (NAIC'96) (pp. 357–368).Utrecht: Utrecht University.

Vreeswijk, G. (1993). *Studies in Defeasible Argumentation*, [Doctoral dissertation, Free University Amsterdam].

Vreeswijk, G. (1997). Abstract argumentation systems. *Artificial Intelligence*, 90, 225–279.

Walton, D. (1984). *Logical dialogue-games and fallacies*. London: Lanham.

Walton, D. (2009). Objections, rebuttals and refutations. *Argument Cultures: Proceedings of OSSA 09*, CD-ROM (pp. 1–10), Windsor, ON: OSSA.

Walton, D. (2011). Modeling burden of proof in a modified Hamblin dialogue system. *Informal Logic*, 31, 279–304.

Walton, D. (2015). Profiles of dialogue: A method of argument fault diagnosis and repair. *Argumentation & Advocacy*, 52, 91–108.

Walton, D., & Gordon, T. F. (2015). Formalizing informal logic. *Informal Logic,* 35(4), 508–538.

Walton, D., & Krabbe, E.C.W. (1995). *Commitment in dialogue*. Albany: SUNY Press.

Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation schemes*. New York: Cambridge University Press.

Wieten, R, Bex, F.J., Prakken, H. & Renooij, S. (2019). Constructing Bayesian network graphs from labelled arguments. In: G. Kern-Isberner and Z. Ognjanovic (eds.), *Proceedings of the 15th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (pp. 99–110) (Springer LNAI Vol. 11726). Cham: Springer.

Yanal, R. J. (1991). *Dependent and independent reasons*. *Informal Logic*, 13, 137–144.

Zenker, F. (Ed.) (2013.) *Bayesian argumentation: The practical side of probability*. Dordrecht: Springer.

Zenker, F. (2015). Denying the antecedent probabilized: a dialectical view. In: Eemeren, F.H. van, Garssen, B., Godden, D., and Mitchell, G. (eds). *Proceedings of the 8th International Conference of the International Society for the Study of Argumentation* (ISSA), July 2014 (pp. 1–14). Amsterdam: Rozenberg.

Zenker, F., Debowska-Kozlowska, K., Godden, D., Selinger, M., Wells, S. (2020). Five approaches to argument strength: probabilistic, dialectical, structural, empirical, and computational. In: Dutilh Novaes, C., Jansen, H., van Laar, J. A. & Verheij, B. (eds.). *Reason to Dissent: Proceedings of the 3rd European Conference on Argumentation, Vol. I* (pp. 653–674) (Studies in Logic: Logic and Argumentation; Vol. 87). London: College Publications.

Zhao. X., Strasser, A., Cappella, J. N., Lerman, C., & Fishbein, M. (2011). A measure of perceived argument strength: Reliability and validity. *Communication Methods & Measures*, 5, 48–75.

Zhao, X., & Cappella, J. N. (2016). Perceived argument strength. In D. K. Kim & J. Dearing (eds.), *Health Communication Research Measures* (pp. 119–126). New York: Peter Lang.