

ARTICLE

Special Issue: Bias in International Law

Pragmatic Inferences and Moral Factors in Treaty Interpretation—Applying Experimental Linguistics to International Law

Benedikt Pirker¹ and Izabela Skoczeń²

¹Law Faculty, University of Fribourg, Fribourg, Switzerland and ²Department of Legal Theory, Jagiellonian University, Cracow, Poland

Corresponding author: benedikt.pirker@unifr.ch; izaskoczen@gmail.com

(Received 27 February 2022; accepted 13 March 2022)

Abstract

The article argues, based on results from massive online survey experiments, that, just as the utterances from ordinary conversation, legal rules can convey a surplus meaning, which is more than just the amalgam of the meanings of the words which are employed in the legal rule's formulation. More precisely, the experiments check whether a typology of the types of this surplus meaning—pragmatic typology—describes adequately the psychological processing of, not only everyday speech, but also legal rules. In two experiments—total N = 733—we find that in morally neutral cases the pragmatic typology adequately describes the psychological processes involved in the interpretation of a legal rule. However, we also find that in morally valenced cases, it is rather the moral inferences carried by participants that shape the pragmatic inferences than the other way around.

Keywords: law and language; international law; treaty interpretation; implicature; experimental linguistics; pragmatics; practical reasoning

This article was first presented at the Workshop on “Behavioural Approaches to International Law”, which was organized as part of the project HRNUDGE funded by the ERC (Grant Agreement 803981, PI Veronika Fikfak) and as part of the ‘Afterlife of Cases’ project funded by the Leids Universiteits Fonds/Dr HA van Beuningen Fonds (www.luf.nl) (PI Daniel Peat). Research for this project has been funded by the Swiss National Science Foundation Spark Grant CRSK-1_190912/1 “International Law, Linguistics and Experimentation (IntLLEx)”. We would like to thank Mira Ariel, Merit Sternau, Anne van Aaken and the participants at the European Society of International Law (ESIL) Interest Group on International Courts and Tribunals’ annual meeting 2020 “Biases and International Adjudication” (Max-Planck-Institute for Procedural Law Luxembourg), the University of Zürich Institute for Public International Law’s Forschungslunch (March 18, 2021), and the ESIL Interest Group on Social Sciences and International Law’s 2021 workshop “Behavioural Approaches in International Law” (University of Hamburg) as well as Jennifer Page, Jack Williams, Noel Struchiner, Ivar Hannikainen and Guilherme Almeida for valuable comments on earlier drafts of this paper and Janice Sumbo for editorial assistance.

This paper is part of Veronika Fikfak’s ‘Human Rights Nudge’ project, which has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No. 803891) and Daniel Peat’s project funded by Leiden University Fund/Dr. H.A. van Beuningen Fonds (www.luf.nl). The workshop *Behavioural Approaches to International Law*, which formed the basis for this Symposium, was held at the Institute for Law and Economics at the University of Hamburg in March 2021 as part of both projects.

A. Introduction

It is often the case that interpreters in international law—often courts or tribunals—are confronted with a treaty text whose ordinary meaning does not contain an explicit solution for the interpretive problem which they face. In these cases, based on the well-known norms on treaty interpretation, such as Article 31 of the Vienna Convention on the Law of Treaties,¹ they rightfully turn to additional means of interpretation and examine e.g. the context of the relevant norm² or the object and purpose of the treaty at issue.³ However, it can be shown that based on state of the art semantic and pragmatic theories,⁴ more precise arguments can be developed on what exactly happens in such cases under the somewhat elusive label of ‘ordinary meaning.’⁵ On this basis, the present article takes up the recent suggestion to apply a modern typology for pragmatic interpretations to international law.⁶ It presents the results of an experiment devised to show the application to legal contexts of the typology developed in recent pragmatic research, namely research explaining how it is possible to convey with a sentence more than just the amalgam of the meanings of the words used. The results from this first experiment open up the possibility to ask an additional question: Is there an influence of moral reasoning on the inferencing of different pragmatic interpretation types? We subsequently present a second experiment providing early evidence that this seems to be the case.

The relevance of our results is twofold. First, and generally, the results show that applying experimental linguistic methods to international law works and can provide a new toolkit to study language-related questions in international law.⁷ More specifically, the results provide evidence that linguistic-pragmatic theories on the level of explicitness of language can be transferred to international law and treaty interpretation. This theoretical basis provides new categories for classifying interpretations of norms with unseen precision with regard to how ‘far-fetched’ these interpretations are.⁸ Second, the possibility emerges to examine systematically by what factors these

¹Out of a rich literature, see, e.g., Brian G. Slocum & J. Wong, *The Vienna Convention and the Ordinary Meaning of International Law*, 46 YALE J. INT’L L. 418 (2021).

²In the recent literature, see, e.g., KATHARINA BERNER, *SUBSEQUENT AGREEMENTS AND SUBSEQUENT PRACTICE IN DOMESTIC COURTS* (2017).

³In this regard, see, e.g., in the recent literature on evolutionary interpretation GEORGES ABI-SAAB, KENNETH KEITH, GABRIELLE MARCEAU & CLEMENT MARQUET, *EVOLUTIONARY INTERPRETATION AND INTERNATIONAL LAW* (2019); Julian Wyatt, *Using Intertemporal Linguistics to Resolve the Problem at the Origin and Core of the Evolutionary Interpretation Debate*, in *EVOLUTIONARY INTERPRETATION AND INTERNATIONAL LAW* (Georges Abi-Saab, Kenneth Keith, Gabrielle Marceau & Clement Marquet eds. 2019); EIRIK BJORGE, *THE EVOLUTIONARY INTERPRETATION OF TREATIES* (2014).

⁴On the usefulness of such theories in international law, see generally Benedikt Pirker & Jennifer Smolka, *Making Interpretation More Explicit: International Law and Pragmatics*, 86 NORDIC J. INT’L L. 228 (2017); Ulf Linderfalk, *What Are the Functions of the General Principles? Good Faith and International Legal Pragmatics*, 78 ZAÖRV 1 (2018); Benedikt Pirker & Jennifer Smolka, *International Law and Linguistics: Pieces of an Interdisciplinary Puzzle*, 11 J. INT’L DISPUTE SETTLEMENT 501 (2020). For more broadly in law, IZABELA SKOCZEŃ, *IMPLICATURES WITHIN LEGAL LANGUAGE* (2019).

⁵On the exact status of ordinary meaning as one means of interpretation or the overall objective of the interpretive process, see Benedikt Pirker, *Balancing Interpretative Arguments in International Law - A Linguistic Appraisal*, 89 NORDIC J. INT’L L. 438, 444 n.26 (2020).

⁶See Jennifer Smolka & Benedikt Pirker, *International Law and Pragmatics: An Account of Interpretation in International Law*, INT’L J. LANGUAGE & L. 5 (2016); Benedikt Pirker & Jennifer Smolka, *Making Interpretation more Explicit: International Law and Pragmatics*, 86 NORDIC J. INT’L L. 1 (2017); Ulf Linderfalk, *Proportionality and International Legal Pragmatics*, 89 NORDIC J. INT’L L. 422 (2020); Jennifer Smolka & Benedikt Pirker, *Pragmatics and the Interpretation of International Law—Two Relevance Theory-Based Approaches*, in *LEGAL MEANINGS: THE MAKING AND USE OF MEANINGS IN LEGAL REASONING* (Janet Giltrow & Frances Olsen eds. 2021); Benedikt Pirker & Jennifer Smolka, *Five Shades of Grey – A Linguistic and Pragmatic Approach to Treaty Interpretation*, ZAÖRV (2022, forthcoming); Izabela Skoczeń, *Is a Legal Implicature Only in the Eye of the Beholder? – On the Interpretation of the CISG Convention*, in *LANGUAGE AND LEGAL INTERPRETATION IN INTERNATIONAL LAW* (Anne-Lise Kjaer & Joanna Lam eds. 2022).

⁷See also Jacob Livingstone Slosser, *Experimental Legal Linguistics: A Research Agenda*, in *LEGAL MEANINGS: THE MAKING AND USE OF MEANINGS IN LEGAL REASONING* (Janet Giltrow & Frances Olsen eds. 2021).

⁸See *supra* note 6.

interpretations seem to be influenced—in our experiment for example moral factors. Thereby, the article contributes to existing debates in (international) law on precision and vagueness in norm-creation and the extent of delegation of authority from law-makers towards legal interpreting agents such as judicial instances that comes with this choice.⁹ The article's main contribution is to provide descriptive tools to better analyze existing legal concepts like ordinary meaning;¹⁰ neither is it the goal to take a position on whether explicitness or vagueness in norms is somehow 'better' nor whether, for example, the means of interpretation of international law are applied correctly.¹¹

In what follows, we first present a linguistic primer on semantics and pragmatics, focusing on the pragmatic typology and the experimental design we developed on its basis. Then, we present the two experimental studies and discuss their results. We end with a general discussion comparing the results of the two experimental studies and with a conclusion on what questions for future experimental research have arisen. Due to certain limitations of the two experimental studies, further experimentation is necessary to fully support our claims.

B. Towards an Experimental Design

1. A Brief Introduction to Linguistics, Semantics and Pragmatics for International Lawyers

Before we turn to the typology of pragmatic interpretations examined in the present article, the basics of the distinction between semantics and pragmatics should be set out.¹² Semantics deals with meaning to the extent that it is *encoded* in the formal components of language. Pragmatics, in contrast, examines meaning as the communication of concepts or thoughts by means of a particular way of *using* such components with encoded meanings in particular contexts.¹³ Semantics uses a *code* model which suggests communication is encoded directly or indirectly in language. According to pragmatics' *inferential* model, the communicator provides evidence of her intention to convey a meaning. In turn, the audience infers meaning based on the evidence provided, the contextual information and their prior knowledge.¹⁴ Take the example of an addressee interpreting the speaker's utterance 'Can you pass me the salt?'. The concept of salt can be decoded; but another part of the meaning has to be inferred from the context, namely whether the utterance is a (polite) request or an actual question. As treaty interpretation in international law relies on ordinary meaning and thus the 'normal' process of human language processing, both the code model and the inferential model are relevant to study interpretation.

To provide a structure to such inferences, linguists like Mira Ariel have developed typologies of pragmatic interpretations. In what follows, we present Ariel's recent, comprehensive effort in this regard.

⁹See, e.g., the debate over rules versus standards in Russell B. Korobkin, *Behavioral Analysis and Legal Form: Rules vs. Standards Revisited*, 79 OR. L. REV. 23 (2000); Gillian K. Hadfield, *Weighing the Value of Vagueness: An Economic Perspective on Precision in the Law*, 82 CALIF. 541 (1994); Kathleen M. Sullivan, *Foreword: The Justices of Rules and Standards*, 106 HARV. L. REV. 22 (1992); Pierre Schlag, *Rules and Standards*, 33 UCLA L. REV. 379 (1985).

¹⁰As far as other means of interpretation like the 'object and purpose' and the 'context' of a treaty also rely on an ordinary meaning approach. RICHARD GARDINER, *TREATY INTERPRETATION* 177, 192 (2010). Our findings also apply to these means of interpretation.

¹¹See, e.g., in recent literature on the question of a hierarchy between these means Donald H. Regan, *Sources of International Trade Law - Understanding What the Vienna Convention Says about Identifying and Using 'Sources for Treaty Interpretation'*, in *THE OXFORD HANDBOOK ON THE SOURCES OF INTERNATIONAL LAW* (Samantha Besson & Jean D'Aspremont eds., 2017).

¹²On the development of the field of pragmatics from its origins as linguistics' 'wastebasket', see Jacob Mey, *How to do Good Things with Words: A Social Pragmatics for Survival*, 4 PRAGMATICS 239, 247 (1993).

¹³Robyn Carston, *Legal Texts and Canons of Construction: A View from Current Pragmatic Theory*, in *LAW AND LANGUAGE* 9 (Michael Freeman & Fiona Smith eds., 2013).

¹⁴Deirdre Wilson & Dan Sperber, *Relevance Theory*, in *THE HANDBOOK OF PRAGMATICS* 607 (Laurence Horn & Gregory Ward eds., 2006); SANDRINE ZUFFEREY & JACQUES MOESCHLER, *INITIATION À L'ÉTUDE DU SENS* 88 (2012).

II. The Core of Mira Ariel's Types of Pragmatic Interpretations

Ariel distinguishes several types of pragmatic interpretations. Ranking the strength of interpretation—which can be understood as the propensity to count as the speaker's relevant contribution—she lists:

- 1) Explicature
- 2) Strong implicature
- 3) Particularized conversational implicatures¹⁵

Based on her own previous research and work by other authors,¹⁶ she suggests that there are a number of tests that can be employed to find out how a particular interpretation is to be categorized. In a nutshell and somewhat simplified for a legal perspective interested mainly in interpretation, Ariel's typology makes it possible to distinguish interpretations on a continuum from literal to ever more far-fetched.

In the following, we explore the typology focusing on the listed levels of meaning that we intend to test in our experimental studies, namely explicatures as well as strong and particularized—'weak'—conversational implicatures, and explain the tests that can be used to identify them.¹⁷ We will additionally rely on 'linguistic meaning' as a level of meaning not present in Ariel's typology. This level consists of simply making the utterance meaning as explicit as possible, as will be shown in our experimental studies below. It serves as a comparative benchmark to the levels of meaning explained below which are to a lesser extent linked to the explicit elements of the utterance at issue. Whereas in the present description the levels of meaning are exemplified with everyday language examples, we will develop legal language examples when presenting the vignettes of our experimental studies.

1. Explicature

In any given situation of interpretation, an addressee must develop certain explicit elements of the utterance at issue. In pragmatics, this is called an explicature. An explicature is indispensable to be able to correctly interpret a speaker's intended meaning.¹⁸ In her article, Ariel generally uses as examples excerpts from a newspaper article on so-called honor killings:

My son said that *she* wasn't the last *one*. We're waiting for the next *one*.

The elements of the utterance in italics must be developed for interpretation, and thus, a resulting explicature would be something like:

The speaker's son said that *Busaina Abu Ghanem* wasn't the last *female murder victim in the family*. We're waiting for the next *female murder victim in the family*.

How can you test whether something is an explicature or not? Most importantly,¹⁹ there is the 'that-is' test. One adds a 'that is (to say)' clause to spell out the explicature. Take our example:

¹⁵Mira Ariel, *Revisiting the typology of pragmatic interpretations*, 13 *INTERCULTURAL PRAGMATICS* 1, 28 (2016). Mira Ariel additionally lists provisional explicature, background assumptions and truth-compatible inferences, but they are not the focus of the present article.

¹⁶For background assumptions, see John Searle, *Literal Meaning*, 13 *ERKENNTNIS* 207 (1978); John Searle, *The Background of Meaning*, in *SPEECH ACT THEORY AND PRAGMATICS* 221 (John Searle, Ferenc Kiefer & Manfred Bierwisch eds., 1980). On implicated premises and conclusions, see DAN SPERBER & DEIRDRE WILSON, *RELEVANCE, COMMUNICATION, AND COGNITION* 44 (2d ed. 1995). On privileged interactional interpretations, see Mira Ariel, *Privileged Interactional Interpretations*, 34 *J. PRAGMATICS* 1003 (2002); KASIA JASZCZOLT, *DEFAULT SEMANTICS: FOUNDATIONS OF A COMPOSITIONAL THEORY OF ACTS OF COMMUNICATION* (2005). On truth-compatible inferences, see Mira Ariel, *Most*, 80 *LANGUAGE* 658 (2004).

¹⁷For a more extensive exploration of Ariel's typology with legal examples, see Jennifer Smolka & Benedikt Pirker, *supra* note 6.

¹⁸MIRA ARIEL, *INTERCULTURAL PRAGMATICS* 4 (2016).

¹⁹As Ariel shows, there is also an additional test, the so-called 'said' test. Because of certain shortcomings of that test (See Herman Cappelen & Ernest Lepore, *On the Alleged Connection Between Indirect Speech and the Theory of Meaning*, 12 *MIND & LANGUAGE* 278 (1997)) and to keep matters simple, we follow here Ariel in simply using the 'that-is' test.

The speaker's son said that she, that is (to say) Busaina Abu Ghanem, wasn't the last one, that is (to say) the last female murder victim in the family. We're waiting for the next one, that is (to say) the next female murder victim in the family.

Explicatures form a single meaning layer with the linguistic meaning; they are developments of the logical form of a sentence.²⁰ The pragmatic inferences are limited to adjustments of the proposition expressed; they are not consciously available as separate interpretations.²¹

An explicature can thus take the form of pronouns that must be pragmatically enriched, e.g., by determining who 'she' is in an utterance. An explicature can also be interpretations of broad or ambiguous terms in a legal norm, as long as the pragmatic interpretation is clearly based on an explicit element of the norm text.

2. Strong Implicature

Strong implicature occurs in contexts where speakers say one thing (a first tier) but intend quite another (a second tier). In the case of strong implicature, a speaker does not express a certain interpretation directly, but instead ultimately intends to replace the directly communicated meaning with the interpretation.²² Ariel's example is the following conversation.²³

R₁: And John Doe, who is a company director, pretends to know that the balance sheet is going to be good so he starts buying.

S: OK that's a criminal offence.

R₂: Eh . . .

S: It's a bit of a criminal offence.

R₃: So he has a mother-in-law.

S: For this you go to jail.

R₃ very strongly implicates in the example that John Doe would illegally buy shares under his mother-in-law's name. This is a strong implicature whose content is separate from what is explicitly stated. Ariel suggests the 'indirect-addition' test as a first step to identify such strong implicature.

The speaker said that John Doe has a mother-in-law, *and in addition he indirectly conveyed that* John Doe would illegally buy shares under his mother-in-law's name. ['indirect addition' test]

To conclusively identify strong implicature, Ariel suggests an extended 'replacement' test—'but actually indirectly conveyed'—that identifies the element of strong implicature and adds the term 'literally'.²⁴

R *literally* said that John Doe has a mother-in-law, *but actually he indirectly conveyed that* John Doe would illegally buy shares under his mother-in-law's name. ['replacement' test]

3. Particularized Conversational Implicature

Contrary to explicature, particularized conversational implicatures (PCIs) are consciously perceived as separate from the content of the explicature on which they are based.²⁵ Take the following example of an utterance by a reporter in Ariel's example newspaper article:

Last Saturday night, Busaina Abu Ghanem was murdered, the *tenth female victim in the family*.

There is a reason why the reporter mentions the fact that it is the tenth female murder victim in the family. Something additional is to be conveyed by the explicit content. This could, for example, be:

²⁰See ARIEL, *supra* note 18, at 11.

²¹*Id.* See also, in turn, *infra* Section 3, Particularized Conversational Implicatures.

²²*Id.* at 4.

²³See *id.* at 19.

²⁴*Id.* at 20.

²⁵This means that they also have their own truth conditions.

There is something terribly wrong with this family.

This interpretation is implicit and can be cancelled explicitly. It is also separate from the relevant explicature. It is different in content and truth conditions: The PCI could not be true without this necessarily affecting whether the main example is true. There could be nothing wrong with the family, or something wrong with this family. This does not change whether Busaina Abu Ghanem was the tenth female victim in the family. As the PCI is different from explicature, if submitted to the ‘that-is’ test, it fails the test. There is, however, a test to identify PCIs, namely again the previously mentioned ‘indirect-addition’ test.²⁶ The mentioned ‘replacement’ test, by contrast, does not work for PCI.

The speaker said that last Saturday night, Busaina Abu Ghanem was murdered, the tenth female victim in the family, *and in addition she indirectly conveyed that there is something terribly wrong with this family.* [‘indirect-addition’ test]

This test targets indirectly communicated messages, for example, implicated conclusions.²⁷ Implicated conclusions depend on contextual assumptions, but the utterance content—the tenth victim—actively participates in shaping the implicated conclusion—there is something wrong.

III. The Experimental Design

As a first step we test whether Ariel’s typology is also a psychological reality in legal interpretation as it would have pertinent consequences for delimiting the notion of ‘ordinary meaning’ employed in treaty interpretation. For the present experimental studies, we rely on laypersons as participants. Our legal justification for proceeding in this way is the ‘ordinary meaning’ rule of treaty interpretation. In international law, ordinary meaning is the result of a comprehension process by a person ‘reasonably informed’ about a treaty; it is not a special understanding that could only be reached by a specialist or international law expert.²⁸ This allows us to draw at least preliminary conclusions from experimental studies with laypersons. Of course, this does not mean that treaty interpretation in international law does not differ at all from interpretation in situations of everyday language communication. But there should be—as international law requires it—some commonalities. While we accept the compromise solution of layperson participation, we nonetheless rely on actual cases in our experimental scenarios to ensure as far as possible a realistic situation of interpretation at least in this regard.

We employ the experimental design by Sternau et al.,²⁹ namely, participants are asked to assess an utterance by an international lawyer about a legal rule, so as to hold the assessed utterance identical across conditions. We employ the same measures as Sternau et al.: truth—a binary measure—confidence about truth judgement and deniability. These measures are based on a theoretical assumption held by Ariel and others:³⁰ linguistic inferences are carried with degrees of strength, depending on the speaker’s commitment to conveyed content and the addressee’s confidence in retrieving the conveyed content.

²⁶MIRA ARIEL, PRAGMATICS AND GRAMMAR 261 (2008).

²⁷One can also draw a distinction between two kinds of PCIs, implicated conclusions and implicated premises (see with further examples from international law Pirker & Smolka, *supra* note 6, at 259. However, Ariel argues that implicated premises are part of background assumptions, and we follow her view in dealing with them under that categorization. ARIEL, *supra* note 18, at 13–14 On the previous definition, see HERBERT PAUL GRICE, STUDIES IN THE WAY OF WORDS 86 (1989).

²⁸See Oliver Dörr, *Article 31 - General rule of interpretation*, in VIENNA CONVENTION ON THE LAW OF TREATIES - A COMMENTARY 581 (Oliver Dörr & Kirsten Schmalenbach eds., 2018), who defines ordinary meaning not as ‘any layman’s understanding’, but as what a person ‘reasonably informed on the subject matter of the treaty’ would understand under the treaty’s terms.

²⁹Merit Sternau, Mira Ariel, Rachel Giora & Ofer Fein, *Levels of interpretation: New Tools for Characterizing Intended Meanings*, 84 J. PRAGMATICS 86 (2015).

³⁰Mainly proponents of Relevance Theory, one of the most prominent approaches in current pragmatics. See, e.g., Deirdre Wilson, *Relevance Theory*, in THE OXFORD HANDBOOK OF PRAGMATICS (Yan Huang ed., 2017); BILLY CLARK, RELEVANCE THEORY (2013).

The deniability measure is a control measure supposed to mirror the results of confidence: the less confident the addressee about an interpretation, the more deniable this interpretation should seem to the addressee.³¹ Deniability is ‘a discursive pragmatic correlate which pertains to the speaker’s perceived ability to actually deny a potential message or inference somehow associated with his/her utterance when in a specific interaction.’³²

As a second step, encouraged by the results of experiment one which confirms that the pragmatic typology applies to treaty interpretation, we introduce an additional level to our investigation in experiment two that we had not originally envisaged, and that arose over the course of our research. Namely, it became clear to us that we can also test an additional claim advocated in jurisprudence theory, but also relevant to international law. This claim is typically argued by the proponents of so-called ‘legal interpretivism,’ which is a legacy of Ronald Dworkin.³³ Modern legal interpretivism claims that moral factors directly influence the perception of legal rules.³⁴ This influence can be governed by different mechanisms than the pragmatic typology. The results of experiment two provide us with relevant results on this claim: there is an influence of moral factors on interpretation, and it affects the psychological processes involved in the pragmatic decoding of meaning.

To sum up, the main question we decided to investigate in our studies is: to what extent does the level of explicitness and moral endorsement of a statement affect the assessment of confidence about truth judgments as well as the deniability of a statement? We formulated three main hypotheses. First, we expected the level of explicitness to influence confidence about truth judgments—higher for higher levels of explicitness. Second, we expected the level of explicitness to influence deniability judgments—lower for higher levels of explicitness.³⁵ Third, we expected the level of moral endorsement of pragmatically conveyed content to influence the patterns described in the first and second hypotheses.³⁶

In this context, there also exists a claim in the literature that it is linguistic pragmatics that influence moral judgments³⁷. We argue in favor of a reverse pattern: It is moral considerations that influence linguistic pragmatic considerations, rather than the other way around. We think that the reason for the direction of this pattern is that the primary function of morality is to guide our own and others’ behavior. Consequently, our moral views shape the way we perceive the world, so that we could direct others at moral goals we find valuable. Thus, moral considerations override our perceptions of meaning described by the pragmatic typology.

C. Experiment 1: ‘Lawfulness’

I. Participants

376 participants were recruited online via the website Amazon Mechanical Turk to perform a survey on the Qualtrics platform. The IP address location of participants was restricted to the

³¹See Sternau et al., *supra* note 29, at 90 (defining deniability as an experimental analogue of Gricean cancelability of implicatures).

³²*Id.*

³³RONALD DWORKIN, *TAKING RIGHTS SERIOUSLY* (1978); Noel Struchiner, Guilherme d’Almeida & Ivar Hannikainen, *An Experimental Guide to Vehicles in the Park*, 15 *JUDGMENT & DECISION MAKING* 1 (2020).

³⁴Marc Greenberg, *Legislation as Communication? Legal Interpretation and the Study of Linguistic Communication*, in *PHILOSOPHICAL FOUNDATIONS OF LANGUAGE IN THE LAW* (2011).

³⁵MERIT STERNAU, *LEVELS OF INTERPRETATION: LINGUISTIC MEANING AND INFERENCES* (2014) (Ph.D. Thesis, University of Tel Aviv).

³⁶See Frank Jackson & Philip Petit, *A Problem for Expressivism*, 58 *ANALYSIS* 239 (1998); Stephen Barker, *Is Value Content a Component of Conventional Implicature?*, 60 *ANALYSIS* 268 (2000); Stephen Finlay, *Value and Implicature*, 5 *PHILOSOPHERS’ IMPRINT* 1 (2005).

³⁷Fred Adams & Annie Steadman, *Intentional Action and Moral Considerations: Still Pragmatic*, 64 *ANALYSIS* 268 (2004); Shaun Nichols & Joseph Ulatowski, *Intuitions and Individual Differences: The Knobe Effect Revisited*, 22 *MIND & LANGUAGE* 346 (2007).

USA. Participants who failed the attention check, took less than a minute to complete the entire survey, or were not native speakers of the English language were excluded, leaving a sample of 270 participants—fifty-two percent of the participants were female; the mean age was thirty-nine years; the age range was nineteen to seventy-nine years.³⁸

II. Methods and Materials

Participants were first presented with an attention check. Thereafter, each participant who passed the attention check was presented with the following information based on a well-known case in European Union law:³⁹

In a previous version of an international organization's founding treaty, it was written that the Court of Justice of the organization could check the lawfulness of acts issued by three of the organization's institutions named the Council, the Commission, and the Parliament. This was due to the fact that all three institutions can influence the legal situation of third parties. Later a new version of the treaty was adopted. The new version of the treaty contains the following legal rule:

Next, participants were split into four groups. Each group received a different, randomly assigned formulation of the legal rule from the vignettes below—the text in bold was removed:

Linguistic meaning The Court of Justice shall review the lawfulness of acts issued by the Council, the Commission, and the Parliament as well as institutions that can influence the situation of third parties.

Explicature The Court of Justice shall review the lawfulness of acts issued by the Council, the Commission as well as institutions that can influence the situation of third parties.

Strong implicature The Court of Justice cannot review the lawfulness of acts issued by institutions that cannot influence the situation of third parties.

Weak implicature The Court of Justice shall review the lawfulness of acts issued by the Council and the Commission.

Later each and every participant from every group saw the exact same information, which was a lawyer's conclusion about the interpretation of the legal rule: "An international lawyer says that the treaty indirectly conveys that the Court of Justice shall review the lawfulness of acts issued by the Parliament." The word 'indirectly' did not appear in the group which was presented with the linguistic meaning vignette.

As can be seen on the basis of the vignettes, we first presented the interpretation by the international lawyer also in the form of the explicit vignette—linguistic meaning—requiring little to no interpretive effort from participants. Then, step by step, we reduced the explicit elements based on Ariel's typology to see whether 'ordinary meaning', in the sense of 'base yourself on the available text for interpretation', would have an effect on participants' reactions to the

³⁸For all the data, appendix and Qualtrics printouts can be found in the online repository, see *Pragmatic Inferences and Moral Factors in Treaty Interpretation—Applying Experimental Linguistic to International Law*, OSFHOME (Nov. 12, 2021), https://osf.io/7529f/?view_only=b96278a6212d484489435db95abba398.

³⁹See Case 294/83, *Parti écologiste 'Les Verts' v. European Parliament*, EU:C:1986:166, paras. 23-25. We did simplify the facts of the case so as to not overburden the laypersons participating in the experiment with information that did not seem useful or relevant to decide the interpretive question at hand. For the present purposes, the distinction between EU law and international law is not relevant, as we make a general case about ordinary meaning interpretation as a legal problem existing in both contexts.

international lawyer's interpretation. To assess participants' reaction, after the vignette, every participant from every group was presented with the exact same question:⁴⁰

Truth Based on the read text, do you think the last sentence is true/false?

Thereafter, all participants were again split into two groups. Each group received a different question. The first group was asked about their confidence about the truth judgment, while the second was asked whether it would be possible to deny the lawyer's conclusion. The exact formulations were (labels in bold omitted):

Confidence How confident are you in your answer? (1 = not at all confident in my answer; 4 = partly confident in my answer; 7 = fully confident in my answer)

Deniability To what extent will the international lawyer be licensed to say in the future: 'In that situation, I did not say that the Court of Justice shall review the lawfulness of acts issued by the Parliament.' (1 = impossible to deny; 7 = completely deniable)

Finally, participants answered a demographics questionnaire.

III. Results

Our result clearly supports the conclusion that Mira Ariel's typology is equally present in legal interpretation as in every-day linguistic interactions. In line with our hypothesis, a decreasing number of participants judged the lawyer's conclusion true with the decreasing level of explicitness of the legal rule. A one-way ANOVA revealed that there was an influence of formulation on the answers on the truth question: $F(3)=28.86$, $p<.001$. Moreover, a chi-square test determined that there was an influence of formulation on the answers on the question on truth: $\chi^2(3)=66.31$, $p<.001$.

Ninety-three percent judged the lawyer's interpretation true when presented with the linguistic meaning of the rule, eighty-nine percent did so with respect to explicature, forty percent with respect to strong implicature and fifty percent with respect to weak implicature. As a side-note, in order to create an example of strong implicature, in the relevant vignette we had introduced a double negation. According to Ariel's typology, double negation can be a strong implicature.⁴¹ Interestingly, the percentages point to the conclusion that double negation should rather be viewed as a weak, rather than strong implicature, because ten percent more participants judged the strong, rather than the weak implicature as false. Thus, perhaps double negation should be avoided in legal text as it is cognitively over-demanding.

⁴⁰Anne van Aaken commented on the phrasing of the following questions that an impersonal phrasing ('would a reasonable person think the last sentence is true/false?' instead of 'do you think the last sentence is true/false?') would have been preferable to reduce the influence of personal biases of participants. While this comment could no longer be taken into account for the present experimental studies, we acknowledge this as a potential factor of influence for our results and will take it into account in the future.

⁴¹For example, take the exchange A: Could humans go to planet Mars? B: This is not impossible. Through the utterance, B conveys that the endeavor is possible, though it is extremely difficult, challenging and has a low probability of success. When applying Mira Ariel's 'replacement' test, one could say that: B literally said that it is not impossible to go to planet Mars but actually he indirectly conveyed that the endeavor is possible, though it is extremely difficult, challenging and has a low probability of success. For more references arguing that double negation can be viewed as implicature, compare Merima Osmankadić, *Why is "Not Infrequent" Not Always "Frequent"? Double Negation in Political Discourse*, 3 EXELL 40 (2015); with Horn, Laurence R. and Heinrich Wansing, 'Negation,' in THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta, ed.).

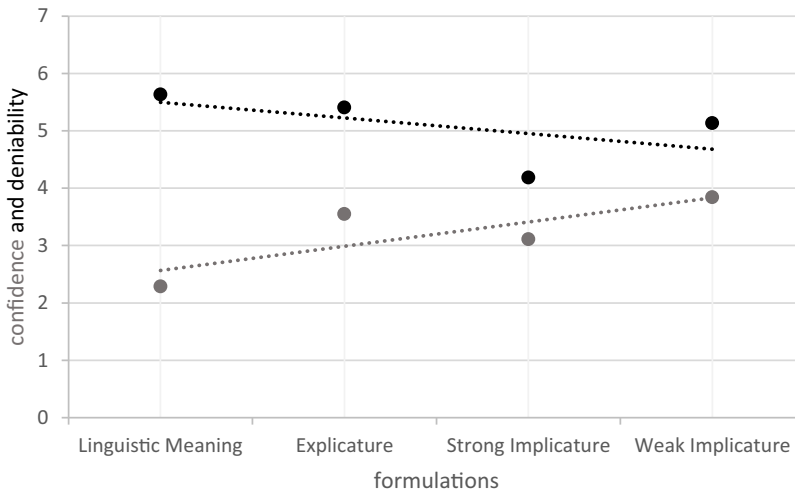


Figure 1. Mean confidence and deniability ratings for each formulation of the legal rule.

Generally, we observed a decrease in reaction times of answering the question on truth between the group who was presented with the linguistic meaning vignette and the remaining vignettes.⁴² It could be the case that the longer reaction time for linguistic meaning was motivated by the fact that participants were somewhat surprised that the rule and interpretation overlapped neatly. We controlled reaction times in order to make sure that none of the conditions would generate reaction times strikingly longer than the others, because this would mean that we are testing a strange case rather than ordinary meaning.

As expected, there was an influence of the formulation of the legal rule on answers to the questions on both confidence and deniability (figure 1). The less explicit the legal rule, the less confident were participants about the lawyer's interpretation. By contrast, the less explicit the legal rule, the more deniable was the lawyer's interpretation.⁴³ Because we replicated neatly the patterns from Sternau et al.'s experiment in every-day language contexts, we conclude that Mira Ariel's typology can be equally applied to legal interpretation.⁴⁴

Looking at the percentages of participants who judge the lawyer's interpretation true and the degrees of confidence judgments, one can notice that there is no difference in inference strength between a 'strong implicature' arising out of double negation and a 'weak implicature' from the real-world legal case used.

IV. Discussion

Linguistic meaning and explicature are assessed most of the time by participants as true sentences, while strong and weak implicature are assessed as false. This provides preliminary ground for a psychological distinction between a proposition derived on the basis of 'what is said' and 'what is implicated'. However, in the tested case, the difference in truth judgments between linguistic meaning and explicature is not pronounced. The weak implicature, which corresponds to the real-world case, generated as many truth as falsity judgments. Thus, we can argue that in such a case interpretation is fully the decision of the court – there seems to be no clear ordinary

⁴²Cf. Appendix table A1.

⁴³Cf. Appendix section 1.

⁴⁴See Sternau et al., *supra* note 29.

meaning for laypersons. There was an influence of formulation (linguistic meaning, explicature, strong implicature, weak implicature) on confidence and deniability judgments.⁴⁵

Finally, some linguistic and legal conclusions can be drawn concerning double-negation—strong implicature.⁴⁶ We observe no difference between the processing of strong and weak implicature, which leads us to a legal conclusion which is important for law drafters: double negation requires a lot of cognitive effort for laypersons and should not be used.

As a general conclusion, our results show that Ariel's typology works also in the context of treaty interpretation. This means that her terminology could be usefully relied upon by international lawyers to discuss difficult cases of interpretation in international law.⁴⁷ Of course, as a crucial caveat, our results are only based on laypersons's answers. Only future studies involving experts can show whether experts' inference patterns are different or similar to laypersons'. At the same time, the fact that the typology seems to work raised an additional question: Could certain behavioral differences be observed in contexts marked by a more morally loaded interpretive task? The 'Lawfulness' setting should appear morally natural for non-expert participants. Legal experts will rather easily understand the constitutional significance of the case with regard to the rule of law and the availability of judicial remedies in the context of an international organization and its organs. Laypersons such as the participants in our experiment, however, arguably would not be aware of such issues, in particular because the way in which we presented the information to them also did not raise the point in a particularly prominent way. The participants thus probably did not hold strong feelings over the case going one way or another. But what if we adopt a different scenario where stronger moral views enter the picture also for non-expert participants? We raised this question in experiment two.

D. Experiment 2: 'Genocide'

Having found that Ariel's pragmatic typology 'works' in legal contexts, we hence decided that additionally it might be interesting to test for the influence of moral attitudes on pragmatic inference. Interpreters in international law are supposed to remain neutral as to personal moral attitudes in the technical exercise of treaty interpretation, namely the identification of ordinary meaning. The lower the explicitness of a treaty norm is, the more an interpreter has to rely on inferences to arrive at a certain interpretation. Still, in all cases an interpreter is to rely on the 'ordinary meaning' to resolve a question of interpretation. If it can be shown that moral attitudes or emotions influence the understanding of 'ordinary meaning' by laypersons, this could indicate that such moral attitudes matter in the formation of inferences. We thus developed an experimental setting based on the *Bosnia Genocide* case⁴⁸ to test this claim. While applying once again Ariel's typology, analogously to experiment one, we examine simultaneously whether

⁴⁵Cf. Appendix experiment 1.

⁴⁶The double-negation in the strong implicature case is mostly regarded as a generalized conversational implicature (GCI). Cf. Horn, Laurence R. & Heinrich Wansing, *supra* note 41; Davis, Wayne, 'Implicature,' in *THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY* (Edward N. Zalta, ed.); IZABELA SKOCZEŃ, *IMPLICATURES WITHIN LEGAL LANGUAGE* (2019). GCI arises in every context in virtue of the words used, namely the two negatives, for instance 'not impossible' means possible (some extreme accounts even see negation as explicature). GCI is distinguished from particularized conversational implicature (PCI), which is highly context dependent. Just as the example above '[I]ast Saturday night, Busaina Abu Ghanem was murdered, the tenth female victim in the family.' Concrete additional contextual information is needed to draw the inference that 'there is something terribly wrong with this family'. GCIs are supposed to be stronger than PCIs. Yet looking at the percentages of participants who judge the lawyer's interpretation true, one can draw several conclusions: (i) double negation is not part of explicature, (ii) there is no difference in strength between a GCI (our strong implicature with double negation) and a PCI (our weak implicature from the real-world case). There is a prominent theory in pragmatics, namely Relevance theory (RT), which claims that GCIs are not necessarily stronger than PCIs, as in RT the strength of implicature depends on cognitive effort rather than semantics/syntax/convention. Consequently, our data points toward RT.

⁴⁷For examples, see also Smolka & Pirker, *supra* note 6.

⁴⁸Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v. Serbia and Montenegro) (Bosnia Genocide), Judgment, 2007 I.C.J. 23, para. 166 (Feb. 26).

the moral attitudes underlying the interpretive exercise have an influence on that exercise. In the present case, these are attitudes concerning the scope of application of the prohibition of genocide as a crime, which, we would argue, is universally recognized as morally abhorrent—both by legal experts and laypersons.

I. Participants

Three-hundred fifty-seven participants were recruited online via the website Amazon Mechanical Turk. The IP address location was again restricted to the USA. Participants who failed the attention check, were not native speakers of the English language or took less than a minute to complete the entire survey were excluded, leaving a sample of 257 participants—fifty-one percent of participants were female; the mean age was forty years; the age range twenty-one to sixty-nine years.

II. Methods and Materials

Having passed the attention check, participants were presented with the following information:

In an international treaty it is written that:

Next, participants were split into four groups. Each group was randomly assigned one of the four vignettes below (labels in bold omitted):

Linguistic meaning The parties agree that genocide is a crime under international law and undertake to prevent, punish and never commit genocide.

Explicature The parties agree that genocide is a crime under international law and undertake to prevent, punish, and never commit it.

Strong Implicature The parties agree that genocide is a crime under international law and undertake to prevent and punish it.

Weak Implicature The parties agree that a serious human rights violation is a crime under international law and undertake to prevent and punish it.

Later, every participant from every group was presented with the same lawyer's interpretation of the legal rule below:

An international lawyer says: 'The treaty states that the parties are under an obligation not to commit genocide.'

After seeing the lawyer's conclusion, all participants had to answer a question on truth assessment of this conclusion:

Truth Based on the read text, do you think the last sentence is: true/false

Participants' reaction times were measured. In a next step, all participants were again split into two groups. One group received a question on confidence about truth judgment, while the other group a question on deniability (labels in bold omitted):

Confidence How confident are you in your answer? (1 = not at all confident in my answer; 4 = partly confident in my answer; 7 = fully confident in my answer)

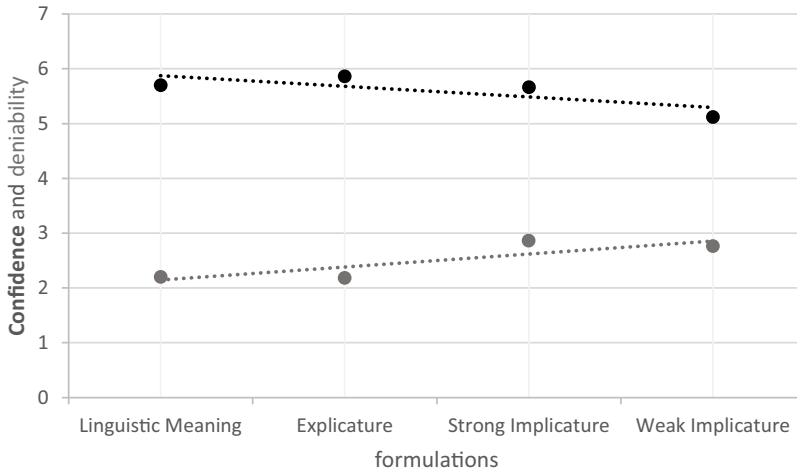


Figure 2. Mean confidence and deniability ratings for each formulation of the legal rule.

Deniability To what extent will the international lawyer be licensed to say in the future: ‘In that situation, I did not say that the treaty states that the parties are under an obligation not to commit genocide.’ (1 = impossible to deny; 7 = completely deniable)

Finally, participants answered a demographics questionnaire.

III. Results

Contrary to experiment one, a similar number of participants judged the lawyer’s conclusion true irrespective of the formulation of the legal rule they were presented with. Namely, ninety-seven percent judged the conclusion true when presented with the linguistic meaning of the legal rule. Ninety-eight percent did so with respect to explicature, eighty-nine percent with respect to strong implicature and eighty-eight percent with respect to weak implicature. Moreover, the reaction times in answering this question were comparable.⁴⁹ A one-way ANOVA determined that there was no influence of formulation on the answer on the truth question: $F(3)=1.49$, $p=.218$. A chi-square test also revealed that there was no influence of formulation on the answers on the truth question: $\chi^2(3)=4.46$, $p=.216$.

There was no influence of formulation of the legal rule on confidence and deniability judgments (figure 2). In other words, irrespective of which formulation of the legal rule participants were presented with, they judged the truth of the lawyer’s interpretation with the same degree of confidence. Moreover, participants judged that it is impossible to deny the lawyer’s conclusion irrespective of the formulation of the legal rule.⁵⁰

IV. Discussion

We find provisionally that based on this example relying on the concept of genocide, moral attitudes have an influence on the willingness to draw inferences in contexts of lower explicitness. There is an influence both on (i) truth and confidence judgments: remaining high despite

⁴⁹Cf. Appendix section 2.

⁵⁰Cf. Appendix section 2.

decreasing explicitness, and on (ii) deniability: remaining low despite decreasing explicitness. We do not even find the trend in mean ratings which was present in experiment one.

An alternative explanation of our findings suggested by a Reviewer was that it is not the influence of morals which distorts inferences described by the pragmatic typology, but rather the fact that the participants knew beforehand that States are under an obligation not to commit genocide, and not (only) that they feel that States should be under such an obligation. There are several factors that could point towards this alternative explanation. The Genocide Convention enjoys broad ratification, meaning that the chances participants have heard of its prohibition under international law is greater. This chance is also greater due to the widespread media coverage of (alleged) incidents of genocide.

We agree that future studies are necessary in order to fully disambiguate between these two hypotheses.⁵¹ However, what could be suggested as potential counterargument, is the fact that if participants knew that committing genocide is prohibited, they also knew that it is prohibited by other legal rules than the ones evaluated in the experiment, e.g., by domestic criminal law. So, participants may just as well have realized that there is no need that the interpretation they are currently evaluating contains a prohibition of genocide, precisely because this prohibition can be found in other legal rules. And yet participants persistently stated that, even if the legal rule does not contain the prohibition, the prohibition is part of its interpretation: an interpretation prohibiting genocide is a true interpretation. In light of this, perhaps the hypothesis of the influence of moral considerations remains—at least partly—valid .

When reviewing a somewhat different experimental design and our findings, Mira Ariel and Merit Sternau commented with regard to a very similar vignette that our chosen phrasing of strong implicature corresponds more to a different category of pragmatic interpretation: a so-called truth-compatible inference. A truth-compatible inference is an inference that ‘the speaker is likely to endorse, given the content of an utterance.’⁵² Moreover, ‘if an assumption turns true in the real world, then it is perceived as compatible with what the speaker said. In other words, the speaker is seen as not having precluded it.’⁵³ Crucially, these inferences are not encompassed by the speaker’s communicative intention, they are merely ‘potentially derived.’⁵⁴ Consider the example: “In the wake of the murder of Busaina Abu Ghanem last weekend, activists say the police must do more to intervene.”⁵⁵

Clearly, it is not the intention of the speaker of the above sentence to convey that there have been complaints filed even before the mentioned complaints by activists, although it is highly likely that the speaker knows there were previous complaints. Rather, the purpose of the sentence is to draw attention to the fact that activists filed a complaint. However, at the same time, the above sentence is compatible with the fact that there were previous complaints.⁵⁶ Mira Ariel thus introduces a ‘compatibility’ test for truth-compatible inferences, here is an example: The journalist said that in the wake of the murder of Busaina Abu Ghanem last weekend, activists say the police must do more. Her utterance is compatible with a state of affairs in which possibly the activists spoke up before—‘compatibility’ test.⁵⁷

In light of this, in ordinary conversation, a strong implicature should be more salient and generate higher truth expectation than a truth-compatible inference. However, if a sentence, or legal rule, from an international treaty is at stake, then we think the distinction between strong implicature and truth-compatible inference becomes blurred both at the psychological—see our results

⁵¹Pirker & Skoczeń, *Inside the Treaty Interpreter’s Mind – An Experimental Linguistics Approach to International Law* (forthcoming 2022).

⁵²ARIEL, *supra* note 18, at 24.

⁵³*Id.*

⁵⁴*Id.*

⁵⁵Smolka & Pirker, *supra* note 6, at 151–52.

⁵⁶*Id.* at 151.

⁵⁷*Id.*

—and theoretical levels. It becomes blurred at the theoretical level because, in the case of a legal rule, it is much less clear who the speaker is and what is her intention. This is because, one could think of many potential speakers having intentions: the signatory States, the representatives of the signatory States, members of parliaments of the signatory States, collective agents such as the parliaments themselves, etc.⁵⁸ Consequently, while in everyday speech it is clear which potential inferences—implicatures—that are compatible with the uttered sentence are chosen by the speaker as intended, the choice is less clear when legal rules are stake. Hence the existence of different interpretive doctrines in common law such as intentionalism, purposivism, textualism etc., and in civil law the linguistic, systemic, and functional doctrines.

We acknowledge the shortcoming of not making completely clear the distinction between strong implicature and truth-compatible inference in our vignette used in experiment two. Therefore, we agree with a reviewer's opinion that our conclusions must encompass the lack of distinction caveat and future studies in this direction are necessary. Nonetheless, as Ariel and Sternau also noted, judging from the results we obtained, participants seem to have been able to infer both strong and weak implicature and seem to not have distinguished between strong implicature and truth-compatible inference in the context of another experiment we run and consulted with Ariel and Sternau. We therefore take this as a lesson for future experimental studies but would argue that our results nonetheless remain at least partly conclusive, even if future studies are necessary.

What is more, our results seem to strengthen the point that morals influence legal interpretation and the deriving of meaning as described by the pragmatic typology. This is because, in experiment two, participants seem to judge a truth-compatible inference just as true as a strong implicature, explicature or as linguistic meaning. This is the case when, we think, strong moral convictions are at play, and thus there is a strong influence of morals on pragmatic considerations. In other words, moral considerations blur the distinction not only between strong implicature and truth-compatible inference, but also between the other listed categories of the pragmatic typology.

E. General Discussion: Comparing Experimental Studies One and Two

In the 'Lawfulness' case tested in experiment one, which we suspected to be morally neutral, the less literal the formulation of the legal rule was, the less participants judged the lawyer's conclusion true. By contrast, in the 'Genocide' case tested in experiment two, which we suspected evoked strong moral feelings,⁵⁹ the number of participants who judged the lawyer's conclusion true was stable independently of the formulation of the legal rule they were presented with (figure 3 and table 1).

The differences between the two cases are also visible in terms of answers to the confidence and deniability questions. In experiment one, Lawfulness case, the less literal the legal rule, the less confident are participants about their truth judgments. By contrast, in experiment two (Genocide case) confidence about truth judgments is stable.

Analogously, in experiment one, according to participants' judgments, the less literal the legal rule, the more deniable the lawyer's conclusion. By contrast, in experiment two, the deniability level is comparable irrespective of the formulation of the legal rule that participants are presented with.

When no strong moral feelings are present, the pragmatic typology is present in participants' answers. However, when strong moral feelings toward the case appear, the pragmatic typology is not visible (figure 4).

Thus, we hypothesize that moral valence has strong influence on truth judgments and the processing of the pragmatic content of an utterance. Participants want the interpretive statement

⁵⁸SKOCZEŃ, *supra* note 46.

⁵⁹See Section D. IV. for our discussion of an alternative explanation suggested by a reviewer.

Table 1. Percentage of participants who judged the lawyer’s conclusion as true for all four formulations—linguistic meaning, explicature, strong implicature and weak implicature—in experiment one, Lawfulness, and experiment two, Genocide

	Formulations			
	Linguistic Meaning	Explicature	Strong Implicature	Weak Implicature
Lawfulness	93%	89%	40%	50%
Genocide	97%	98%	89%	88%

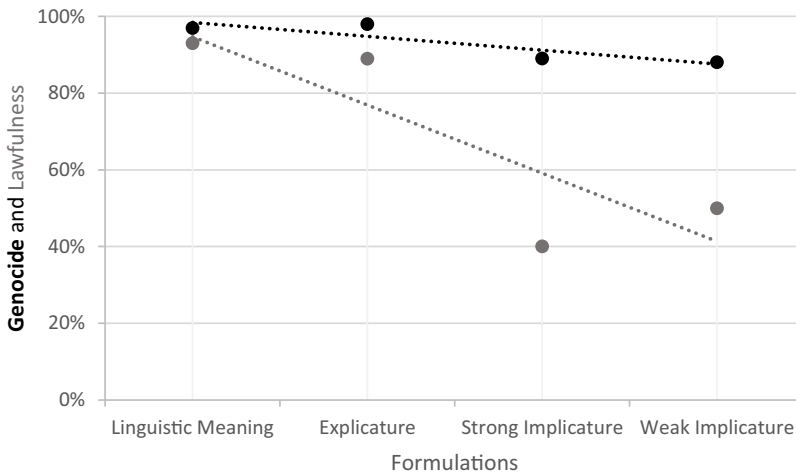
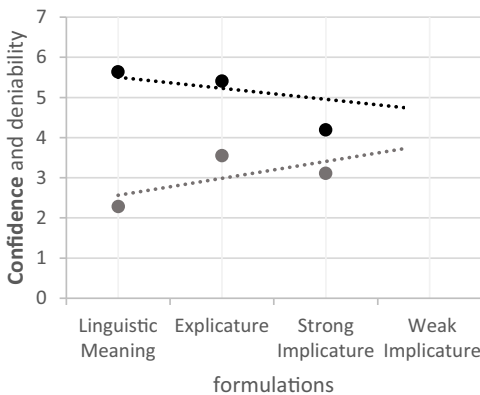


Figure 3. Percentage of participants who judged the lawyer’s conclusion as true for all four formulations—linguistic meaning, explicature, strong implicature, and weak implicature—in the ‘Genocide’ and ‘Lawfulness’ cases.

Lawfulness case



Genocide case

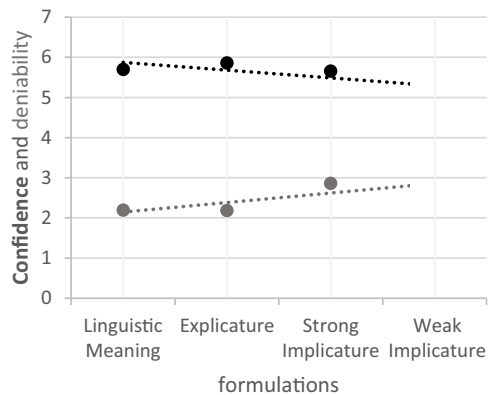


Figure 4. Mean confidence about truth judgment and deniability ratings for experiment 1—Lawfulness case on the left—as well as mean confidence about truth judgment and deniability ratings for experiment 2—Genocide case on the right.

to be true because it is in line with their moral views. Thus, participants rate the interpretation true even if it has no grounds in the interpreted legal rule or the psychological reality of the pragmatic typology.

At the same time, with regard to the mentioned debate on whether linguistic pragmatics influences moral judgments or vice versa,⁶⁰ our results point to moral considerations influencing linguistic pragmatic considerations, rather than the other way around.

As a caveat, we should note that in order to ascertain that our hypothesis on the influence of moral attitudes on legal interpretation is well-founded, future experimental studies will need to be carried out with an additional question inquiring about the level of moral valence of the tested scenario. Moreover, the two cases we tested in the two experimental studies are quite different. We cannot exclude with certainty that topic differences and the differences in availability of the vignettes' subjects in the widely accessible media have introduced 'noise' and have influenced the results limiting their comparability. In future research, a clear goal will be to limit such noise by developing an experimental design which will present only minimal differences between the morally neutral and morally valenced conditions.

F. Conclusion

The present article has suggested that, following up on existing theoretical claims,⁶¹ pragmatics and the presented typology of pragmatic interpretations can help international lawyers sharpen their understanding of language meaning in interpretive processes. The article supported its claims through experimental studies. We have demonstrated that Ariel's typology and its tests can be applied to categorize interpretations with the help of international law examples. Second, in recent years pragmatics has increasingly been relying on experimental approaches.⁶² As international law scholarship catches up in this regard,⁶³ the present paper contributes to this trend by implementing experimental studies based on Ariel's and Sternau's earlier research.

Our findings support the claim that even in legal interpretation the pragmatic typology is a psychological reality. Moreover, moral content has influence on the pragmatic processing of laypersons. Our experimental studies nonetheless leave us with certain additional questions to be explored in future experimental studies. Most notably, similar experimental studies should be conducted to assess how international law experts perform as compared to laypersons, both with regard to morally neutral and morally non-neutral cases of interpretation. In experimental settings, the comparability of the results between morally neutral and morally valenced scenarios should be improved. The issue of whether an interpretive issue is morally neutral could additionally be included into the experimental design as a question to be answered directly by the

⁶⁰Fred Adams et al., *supra* note 37; Shaun Nichols & Joseph Ulatowski, *supra* note 37; Julia Driver, *Attributions of Causation and Moral Responsibility*, in *MORAL PSYCHOLOGY VOLUME 2: THE COGNITIVE SCIENCE OF MORALITY: INTUITION AND DIVERSITY* at 423–29; Noel Struchiner, Ivar R. Hannikainen & Guilherme da F.C.F. de Almeida, *An Experimental Guide to Vehicles in the Park*, 15 *JUDGEMENT & DECISION MAKING* 312 (2020); Brian Flanagan, Guilherme da Franca Couto Fernandes de Almeida, Noel Struchiner & Ivan R. Hannikainen, *Moral Appraisals Guide Intuitive Legal Determinations*, (2021), https://www.researchgate.net/publication/355844887_Moral_Appraisals_Guide_Intuitive_Legal_Determination.

⁶¹Smolka & Pirker, *supra* note 6.

⁶²See IRA NOVECK, *EXPERIMENTAL PRAGMATICS: THE MAKING OF A COGNITIVE SCIENCE* (2018).

⁶³See Adam Chilton & Dustin Tingley, *Why the Study of International Law Needs Experiments*, 52 *COLUM. J. TRANSNAT'L L.* 173 (2013); Yahli Shereshevsky & Tom Noah, *Does Exposure to Preparatory Work Affect Treaty Interpretation? An Experimental Study on International Law Students and Experts*, 28 *EJIL* 1287 (2018); Jeffrey Dunoff & Mark Pollack, *Experimenting with International Law*, 28 *EJIL* 1317 (2017); Anne Van Aaken, *Experimental Insights for International Legal Theory*, 30 *EJIL* 1237 (2019); Tomer Brodeur, *Outcome Bias and Expertise in Investigations under International Humanitarian Law*, 30 *EJIL* 1303 (2019); Anton Strezhnev, Beth A. Simmons & Matthew D. Kim, *Rulers or Rules? International Law, Elite Cues and Public Opinion*, 30 *EJIL* 1281 (2019); Daniel Statman, Raanan Sulitzeanu-Kenan, Micha Mandel, Michael Skerker & Stephen De Wijze, *Unreliable Protection: An Experimental Study of Experts' In Bello Proportionality Decisions*, 31 *EJIL* 429 (2020).

participants. Lastly, in terms of the results of such an inquiry, we could receive hints as to whether international law experts are similarly guided by their moral convictions like laypersons, or rather, whether legal training renders them immune to moral, extra-legal considerations. At the same time, certain conclusions on the effect of the norms of interpretation of international law should also become possible, based on a comparison between the results of laypersons and experts, given a similar interpretive scenario.

Appendix

Experiment 1

Results for the question on truth

Table A1. Percentage of participants that judged the lawyer's utterance as true as well as their reaction times in milliseconds for answering this question for experiment one

	Linguistic Meaning	Explicature	Strong Implicature	Weak Implicature
Truth	93%	89%	40%	50%
Reaction Time	50.20	36.80	42.05	35.31

Results for the questions on confidence and deniability

There was an influence of formulation on confidence and deniability respectively: one way ANOVA for confidence $F(3)=6.54$, $p<.001$; for deniability $F(3)=3.57$, $p=.016$.

Table A2. Mean ratings for confidence and deniability for experiment 1, standard deviations are given in brackets

	Confidence that true	Deniability
Linguistic Meaning	5.64 (1.25)	2.34 (1.53)
Explicature	5.41 (1.10)	3.64 (1.82)
Strong Implicature	4.19 (1.28)	3.17 (1.71)
Weak Implicature	5.13 (1.36)	3.28 (1.80)

Table A3. Independent samples t-test comparing different formulations (LM = linguistic meaning, E = explicature, SE = strong implicature, WI = weak implicature). Confidence intervals are given for the means

	Confidence that true				Deniability			
	<i>t(df)</i>	<i>p</i>	<i>Cohen's d</i>	<i>95% CI</i>	<i>t(df)</i>	<i>p</i>	<i>Cohen's d</i>	<i>95% CI</i>
LM v E	.79(63)	.434	.20	[-.35;.81]	-3.09(60)	.003	.79	[-2.07;-.44]
E v SI	3.42(46)	.001	1.05	[-.83;1.71]	.70(38)	.491	.27	[-.83;1.71]
SI v WI	-2.00(29)	.055	.72	[-1.91;.02]	-1.06(26)	.299	.43	[-2.15;.69]
LM v WI	1.57(68)	.120	.36	[-.129;1.089]	-2.30(65)	.024	.53	[-1.75;-.13]

Experiment 2

Results for the question on truth

Table A4. Percentage of participants that judged the lawyer's utterance as true as well as their reaction times in milliseconds for answering this question for experiment two

	Linguistic Meaning	Explicature	Strong Implicature	Weak Implicature
Truth	97%	98%	89%	88%
Reaction Time	25.24	28.57	27.11	26.68

Results for the questions on confidence and deniability

There was no influence of formulation on confidence and deniability respectively: one way ANOVA for confidence $F(3)=1.34$, $p=.266$; for deniability $F(3)=1.78$, $p=.155$.

Table A5. Mean ratings for confidence and deniability for experiment 2, standard deviations are given in brackets

	Confidence that true	Deniability
Linguistic Meaning	5.69 (1.42)	2.20 (1.41)
Explicature	5.77 (1.53)	2.18 (1.64)
Strong Implicature	5.76 (1.15)	2.86 (1.96)
Weak Implicature	5.26 (1.76)	2.76 (1.28)

Table A6. Independent samples t-test comparing different formulations (LM = linguistic meaning, E = explicature, SE = strong implicature, WI = weak implicature). Confidence intervals are given for the means

	Confidence that true				Deniability			
	<i>t(df)</i>	<i>p</i>	<i>Cohen's d</i>	<i>95% CI</i>	<i>t(df)</i>	<i>p</i>	<i>Cohen's d</i>	<i>95% CI</i>
LM v E	-.20(53)	.842	.05	[-.88;.72]	.06(67)	.949	.01	[-.71;.76]
E v SI	.03(53)	.977	.01	[-.72;.74]	-1.56(67)	.123	.42	[-1.55;.19]
SI v WI	1.23(50)	.226	.34	[-.32;1.31]	.23(67)	.818	.06	[-.71;.89]
LM v WI	.97(50)	.335	.29	[-.46;1.31]	-1.74(67)	.086	.35	[-1.21;.08]