doi:10.1017/aju.2021.34

SYMPOSIUM ON LIMITATIONS OF THE BEHAVIORAL TURN IN INTERNATIONAL LAW

THE COGNITIVE PSYCHOLOGY OF RULES OF INTERPRETATION IN INTERNATIONAL LAW

*Anne van Aaken**

While Articles 31 and 32 of the Vienna Convention on the Law of Treaties (VCLT) prescribe the rules of interpretation for international treaty law as "disciplining rules,"[1] the rules of interpretation themselves are understudied from a cognitive psychology perspective. This is problematic because, as Jerome Frank observed, "judges are incurably human,"[2] like everybody else. I submit that behavioral approaches could provide insights into how biases and heuristics affect the way judges and other interpreters use the VCLT rules.

Hermeneutics, understood as the theory of interpretation, investigates some questions that are also asked in the cognitive sciences. The nature of human understanding, the way that we gain and organize knowledge, the role played by language,[3] and the relations between conscious and unconscious knowledge are all examples from the intersection of hermeneutics and the cognitive sciences.[4] Research in cognitive psychology has demonstrated that human (and judicial) responses deviate from the rationality suggested by various models of decision-making and rational judgment.[5]

What are the implications for our understanding of the rules of interpretation in international law and their disciplining potential? Are some rules of interpretation more prone to biases and heuristics than others and do those biases and heuristics differ for the different rules of interpretation? Research has concentrated on showing experimentally that adjudicators fall prey to biases in *substantive* decisions, focusing on experimental variations in substantive law or facts. They have not looked at the *second order rules of interpretation* and how their use may be influenced by cognitive biases. I proceed as follows: after briefly contrasting the psychological approach with rationalist reasoning, I turn to different methods of interpretation and discuss how behavioral assumptions play a role therein, before discussing the limitations of this endeavor. I draw on nascent available empirical research but much of the remainder remains speculative.

* *Alexander von Humboldt Professor, Chair for Law and Economics, Legal Theory, Public International Law and European Law, Director, Institute of Law and Economics, Universität Hamburg Faculty of Law, Hamburg, Germany.*

[1] *See* Owen Fiss, *Objectivity and Interpretation*, 34 STAN. L. REV. 739, 744 (1982): "He [the interpreter] is disciplined by a set of rules that specify the relevance and weight to be assigned to the material (e.g., words, history, intention, consequence)."

[2] Jerome Frank, *Are Judges Human? Part One: The Effect of Legal Thinking of the Assumption that Judges Behave Like Human Beings*, 80 U. PA. L. REV. 17, 24 (1931).

[3] For constructivist scholars, the ability of language to frame thinking has always been prominent; Nicholas Onuf, *Do Rules Say What They Do - from Ordinary Language to International Law*, 26 HARV. INT'L L.J. 385 (1985), using speech act theory.

[4] Shaun Gallagher, *Hermeneutics and the Cognitive Sciences*, 11 J. CONSCIOUSNESS STUD. 162 (2004).

[5] Keith Stanovich & Richard West, *Individual Differences in Reasoning: Implications for the Rationality Debate?*, 23 BEH. BRAIN SC. 645 (2000).

Cognitive Psychology

The rationality assumption in adjudication, whereby interpreters decide objectively and without cognitive errors, has long been challenged on numerous accounts, both theoretically and empirically. Legal theorists have often articulated their suspicion that interpretation is the result of the result—interpreters tend to confirm their initial intuition by interpretation. Cognitive studies can help to illuminate that suspicion and develop counter-strategies.

Experiments on the psychology of judging[6] demonstrate that cognitive biases and heuristics affect the decision-making of national court judges[7] and international arbitrators.[8] Along with implicit biases (stereotype and attitude-based), those experiments show bounded rationality entering legal interpretation. Biases and heuristics are both associated with dual-process theories,[9] whereby System 1 decisions are made quickly and almost automatically, sometimes leading to mistakes, while System 2 decisions are made deliberatively, connected to logical thinking and reasoning, and less prone to cognitive errors. Ideally, legal interpreters always reason via System 2.

Experimental evidence supports an intuitive-override model of adjudication which posits that judges generally make intuitive decisions but sometimes override their intuition with deliberation.[10] The model thus rejects a pure formalist model, which suggests that judges apply the law logically and mechanically, as well as a pure realist model, whereby judges decide by feelings and hunches.

Psychological Factors in the Use of Rules of Interpretation of the VCLT

Rules of interpretation are meant to help the interpreters in overriding intuition. But they may be prone to biases and heuristics themselves. VCLT Articles 31 and 32 mandate the rules of treaty interpretation to be used, with ordinary meaning, context, and object and purpose being the starting point.[11]

Ordinary meaning focuses on the specific words chosen. But wording is subject to the framing effect.[12] As Shiri Krebs observes, "[l]egal terms . . . create 'frames' under which information is categorized and interpreted. Different legal terms therefore have a distinct impact on people's cognition."[13] In short: words matter, including in legal texts. Whether something is labelled a "war crime"—an emotionally charged term—or a "violation of international law"[14] has consequences for the interpretation, triggering emotional biases. Exploring the choice of particular words in legal texts, their connotations, and their invocation of biases is underresearched and worth exploring through experiments, including experimental pragmatics, a nascent field of research.[15]

[6] *See generally* THE PSYCHOLOGY OF JUDICIAL DECISION MAKING (David E. Klein & Gregory Mitchell eds., 2010).

[7] *See* Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 777 (2001); Chris Guthrie et al., *Blinking on the Bench: How Judges Decide Cases*, 93 CORNELL L. REV. 1 (2007).

[8] Susan Franck et al., *Inside the Arbitrator's Mind*, 66 EMORY L.J. 1115 (2017).

[9] DANIEL KAHNEMAN, THINKING, FAST AND SLOW (2011).

[10] Guthrie et al., *Judicial Mind*, *supra* note 7.

[11] As stated early on in I OPPENHEIM'S INTERNATIONAL LAW 1267 (Robert Jennings & Arthur Watts eds., 9th ed. 1992): "The finding whether a treaty is clear or not is not the starting point but the result of the process of interpretation."

[12] Daniel Kahneman & Amos Tversky, *Choices, Values, and Frames*, 39 AM. PSYCHOLOGIST 341 (1984).

[13] Shiri Krebs, *Law Wars: Experimental Data on the Impact of Legal Labels on Wartime Event Beliefs*, 11 HARV. NAT'L SEC. J. 106, 122 (2020) with further references.

[14] *Id.*

[15] Benedikt Pirker & Jennifer Smolka, *Making Interpretation More Explicit: International Law and Pragmatics*, 86 NORDIC J. INT'L. L. 228 (2017).

https://doi.org/10.1017/aju.2021.34 Published online by Cambridge University Press

Context is closely connected to ordinary meaning. The main reason for looking to the context is to confirm an ordinary meaning. Yet, context can include very different factors ranging from surrounding words, heading of articles, and punctuation, to more remote elements such as other provisions in the same text, similar provisions in other treaties, or even the object and purpose.[16] There is thus a wide range of context elements open to choose from. Two cognitive biases may kick in here. First, the availability bias is a mental shortcut that relies on immediate examples or clues that come to a given person's mind when evaluating a specific topic, concept, method, or decision. A specific context may be made salient by counsel, by previous decisions, or in deliberations, leading to the neglect of other elements. Second, the cognitive bias of motivated reasoning is well known.[17] Motivated reasoning refers to people's tendencies to conform their assessments of information, including from logical arguments or their own sense impressions, to some end or goal extrinsic to judgment accuracy; in other words, motivated reasoning is the "tendency to find arguments in favor of conclusions we want to believe to be stronger than arguments for conclusions we do not want to believe."[18] This is not to say that motivated reasoning necessarily occurs when choosing the context, but it suggests that when using the wide range of possible context elements, the interpreter should be aware of these cognitive biases and search for and deliberate on counter contexts.

Another VCLT requirement is to assess a treaty in the light of its object and purpose, linked to the principle of effectiveness.[19] Since the object and purpose rule requires instrumental rationality (the respective interpretation should effectively achieve the object and purpose), the interpreter is asked to consider real world consequences. This suggests the utility of integrating social science approaches, which have methodologies to illuminate exactly those consequences, in the interpretive process. Finding the object and the purpose of a treaty is not an easy task. The object and purpose can be found in the preamble or sometimes even the title of the treaty.[20] But often it seems that it is found by the courts by intuition without transparent explanation.[21] Several questions therefore arise for an interpreter: what is the meaning of the object and purpose, how are they to be identified, what is the relationship between the object and the purpose, how can multiple purposes of a treaty be balanced, and, I propose, whether and how social science insights should be taken into account when giving effect to the object and the purpose.

Consider the example of international investment law. Often there are several, possibly conflicting, purposes,[22] but many tribunals have considered only investment protection, without considering (sustainable) development or "the welfare of both peoples"—principles also found in preambles of the relevant treaties. Why have the tribunals been so fixated on investor protection? A possible explanation may be anchoring. This is a cognitive bias whereby an individual's decisions are influenced by a particular reference point or "anchor."[23] For example, if someone were to see a pair of shoes priced at US$500 and then later to encounter a pair for US$200, they would find the latter pair to be cheap. Legal anchoring may also play a role in the choice of focus for the object and purpose.

[16] Richard K. Gardiner, Treaty Interpretation 178 (2008).

[17] Dan M. Kahan, *Ideology, Motivated Reasoning, and Cognitive Reflection*, 8 Judgment & Decision Making 407 (2013).

[18] Ziva Kunda, *The Case for Motivated Reasoning*, 108 Psychol. Bull. 480, 480–81 (1990).

[19] Isabella Van Damme, *Treaty Interpretation by the WTO Appellate Body*, 21 EJIL 605, 630 (2010).

[20] Certain Norwegian Loans (Fr. v. Nor.), 1957 ICJ Rep. 9, 24 (July 6).

[21] Isabelle Buffard & Karl Zemanek, The "Object and Purpose" of a Treaty: An Enigma?, 3 Austrian Rev. Int'l & Eur. L. 311, 329 (1998).

[22] David S. Jonas & Thomas N. Saunders, *The Object and Purpose of a Treaty: Three Interpretive Methods*, 43 Vand. J. Transnat'l L. 565 (2010).

[23] The notion of anchoring has been addressed in the general legal literature in several contexts. Potentially most important, studies have found that the interpretation of vague norms can be influenced by legal anchoring. Yuval Feldman et al., *Anchoring Legal Standards*, 13 J. Emp. Leg. Studies 298 (2016).

https://doi.org/10.1017/aju.2021.34 Published online by Cambridge University Press

In investment law, early prominent cases used "investment protection" as the sole object and purpose[24] and this may have anchored subsequent tribunals, obscuring or diminishing other options. If "welfare" or (sustainable) development—as also stated in many preambles—had instead been used as the relevant object and purpose, tribunals would have had to address the question whether extensive investment protection contributes to development or welfare. That is a social science question which has been empirically explored. This research shows that a direct causal relationship between strong investment protection and welfare or development is by no means secured and the judicial intuition that it exists may be wrong.[25]

The relationship between VCLT Article 31 and the supplementary rules of interpretation of Article 32 remains debated.[26] The use of *travaux préparatoires* is one of the elements of Article 32 which has been experimentally researched. Many studies have demonstrated that it is difficult if not impossible for juries and judges[27] to ignore inadmissible evidence even under instructions to ignore such information, following a known bias against disregarding even false information. One path-breaking international law study explores the ability to ignore the *travaux préparatoires* of a treaty provision in the process of interpretation.[28] It uses experiments administered to both laypersons (law students) and experts (international law practitioners and academics). It is unusual in that it shows that experts are more immune to cognitive biases than laypersons and that the latter are not fully immune to such cognitive bias even if they state that they are aware of the diminished legal status of *travaux*.

Beyond the VCLT, other interpretative practices may be subject to biases as well. Proportionality analysis, pervasive in international law, is a mode of reasoning, a methodological tool used for balancing competing rights or principles,[29] be it in human rights law, trade law, or international humanitarian law. Hitherto, it has rarely been tested experimentally,[30] but it is highly likely that its decision-architecture and how the ingredients of proportionality analysis are presented may produce biases (or counter them).[31] Different steps of proportionality analysis, namely analyzing suitability and necessity depend on an assessment of probabilities. The steps of suitability and necessity necessarily incorporate a weighing of the probability that the measure(s) chosen will achieve the purpose. To give an example: is a ban on civil servants joining unions a suitable and necessary (least restrictive) means to protect the functioning of the state?[32] Cognitive psychology has interesting insights concerning the assessment of probabilities. To name just two: the base rate neglect, that is, the tendency to base judgments on specifics, ignoring general statistical information, as well as the probability neglect, that is, the tendency to completely disregard or

[24] Amco Asia Corporation v. Indon., ICSID Case No ARB/81/1, Decision on Jurisdiction, para 23 (Sept. 25, 1983), 23 ILM 351 (1984); Siemens v. Arg., ICSID Case No ARB/02/8, Decision on Jurisdiction, para 81 (Aug. 3, 2004).

[25] Anne van Aaken & Tobias Lehmann, *Sustainable Development and International Investment Law: An Harmonious View from Economics*, *in* Prospects in International Investment Law and Policy 317 (Roberto Echandi & Pierre Sauvé eds., 2013).

[26] Julian D. Mortenson, *The Travaux of Travaux: Is the Vienna Convention Hostile to Drafting History?*, 107 AJIL 780 (2013).

[27] Andrew J. Wistrich et al., *Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding*, 153 U. Pa. L. Rev. 1251 (2005). For other areas of judgement, see Moti Mizrahi, *Arguments from Expert Opinion and Persistent Bias*, 78 OSSA Conf. Archive 1 (2016).

[28] Yahli Shereshevsky & Tom Noah, *Does Exposure to Preparatory Work Affect Treaty Interpretation? An Experimental Study on International Law Students and Experts*, 28 EJIL 1287 (2017).

[29] Ahron Barak, Proportionality 131 (2012).

[30] One study tests proportionality by conducting a survey with legal experts in Israel concerning proportionality in targeted killing scenarios: Raanan Sulitzeanu-Kenan et al., *Facts, Preferences, and Doctrine: An Empirical Analysis of Proportionality Judgment*, 50 Law & Soc'y 348 (2016).

[31] For a detailed analysis of how biases can influence proportionality analysis, see Anne van Aaken, *The Decision Architecture of Proportionality Analysis: Cognitive Biases and Heuristics* (2019).

[32] Demir and Baykara v. Turk., App. No. 34503/97 (Eur. Ct. H.R. 2018), *see also* Ragnar Nordeide, *Demir & Baykara*, 103 AJIL 567 (2009).

https://doi.org/10.1017/aju.2021.34 Published online by Cambridge University Press

overestimate probability when making a decision under uncertainty with small probabilities. Both can lead to distorted assessments of probabilities in the proportionality assessment.

A further interpretive practice is used by the European Court of Human Rights: the "European Consensus." In an interesting study, Daniel Peat draws on a data set of 461 judgments of the Grand Chamber of the Court and suggests that the Court's use of said doctrine must be understood partially as an aversion to choice overload, that is, the cognitive difficulty many people have in making decisions when faced with many options. Psychological reasons may thus better explain the use of the practice than reference to the VCLT rules.[33]

Limitations to Using Behavioral Approaches for Understanding Rules of Interpretation

Many challenges and limitations remain. First, many experiments have been conducted with national judges; it is basically impossible to conduct experiments with international judges. Even if some of them would consent to participate, the number necessary to conduct meaningful research is likely to be too small. Although many studies have found cognitive biases in expert adjudicators, there have also been studies showing that those may be diminished in experts compared to other populations. There are thus limits to experiments conducted only with laypersons such as students, impacting external validity, that is the generalizability of the findings to other situations and people.

Second, given those difficulties, using text as data or hand-coding decisions of international courts and tribunals, as done in the study by Peat, is another way to go, but this type of research method does not have the same internal validity[34] as experiments, since causal factors are more difficult to disentangle. Thus, one of the main challenges in applying behavioral insights to interpretation via data analysis is working out what indicators suggest that an interpretation is influenced by psychological cognitive biases as opposed to other, e.g., contextual, factors. It is crucially important to identify relevant independent variables (explanatory variables) for quantitative analysis.

Translating experimental insights into normative guidance has to be done very carefully, and only after reliable empirical evidence, which is still missing. Furthermore, social science methodology is complex and requires expertise not generally held by lawyers, so translation between the fields may pose difficulties.

Conclusion

Shedding light on how psychology may influence how judges or other officials use the rules of interpretation in international law is not meant to delegitimize legal interpretation by international courts or other interpreters. Rather, it is meant to help to mitigate the so-called blind spot bias (the bias of biases) in legal interpretation, namely that people tend to ignore their own cognitive biases even as they see them in others. Becoming aware of those potential biases can help the interpreter to overcome them and use the rules of interpretation consciously and transparently. Also, institutional arrangements, like the composition of decision-making groups, can help mitigate them. By sharpening our understanding of how rules and practices of interpretation actually operate, the normative debate can evolve in new directions.

[33] Daniel Peat, *The Tyranny of Choice and the Interpretation of Standards: Why the ECtHR Uses Consensus*, NYU J. Int'l L. & Pol. 381 (2021).

[34] Internal validity describes a trustworthy cause-and-effect relationship between a treatment and an outcome.

https://doi.org/10.1017/aju.2021.34 Published online by Cambridge University Press