The European Journal of International Law Vol. 28 no. 4

Does Exposure to Preparatory Work Affect Treaty Interpretation? An Experimental Study on International Law Students and Experts

Yahli Shereshevsky* and Tom Noah**

Abstract

Ample research has demonstrated that exposure to inadmissible evidence affects decision making in criminal and civil cases. However, the difficulty of ignoring information in the context of legal interpretation has not been examined yet. Our study addresses the possible effects that exposure to preparatory work has on the interpretation of treaties. In the present article, we examine the ability of students enrolled in international law courses and of international law experts to ignore preparatory work when they are not allowed to use it. We found that exposure to preparatory work affected the students' interpretation of treaties, while no such effect was found among the experts. These results reaffirm the practical relevance of the debate over the hierarchy between the rules of treaty interpretation. In particular, our study demonstrates that preparatory work can play a significant role in decision making, depending on the legal rule that applies to the use of such materials. More generally, our study suggests that legal interpretation by students and experts is qualitatively different and that international law experts might be better able than non-experts to discount irrelevant information in the process of treaty interpretation.

* Michigan Grotius Research Scholar, University of Michigan Law School, Ann Arbor, MI, USA. The research was conducted when the author was a Hans Guth Dreyfus Doctoral Fellow at the Hebrew University of Jerusalem Law School. Email: yahli.shereshevsky@gmail.com.

** PhD Candidate, Department of Psychology, Hebrew University of Jerusalem, Jerusalem, Israel. Email: tom.noah@mail.huji.ac.il.

EJIL (2017), Vol. 28 No. 4, 1287–1316 doi:10.1093/ejil/chx069

1 Introduction

Does exposure to the preparatory work of a treaty inevitably affect its interpretation even under a rule that prohibits its use? Are legal experts any better able than students to ignore its influence? This article engages in the ongoing debate over the role of preparatory work under the Vienna Convention of the Law of Treaties' rules of treaty interpretation (VCLT rules) by exploring these questions empirically.[1]

One of the major debates over treaty interpretation concentrates on the possible hierarchy between the VCLT rules.[2] Articles 31 and 32 of the VCLT set out, respectively, the 'General Rule' and the 'Supplementary Means' of treaty interpretation. According to the traditional approach, these rules establish a strict hierarchy between Articles 31 and 32 of the convention. The interpreter must begin the interpretation by applying the general rule, which includes, *inter alia*, textual, contextual and teleological methods of interpretation. Recourse to the supplementary means of interpretation, which include, *inter alia*, the preparatory work of the treaty, is allowed only under limited conditions. It is always permissible to use a supplementary means of interpretation in order to confirm a meaning that results from the application of the general rule, but, in order to determine the meaning of a text, the supplementary means of interpretation may be used only if application of the general rule leaves that meaning 'ambiguous or obscure; or lead[s] to a result which is manifestly absurd or unreasonable'. This traditional approach has been challenged in recent years by several authors who argue that the VCLT rules do not establish a strict hierarchy between the general rule and the supplementary means of interpretation (the corrective approach). What is at

Earlier versions of the article were presented at the 2016 ASIL Research Forum, the American Law and Economics Association 2016 annual meeting, the 2016 Impact of Law on Behaviour conference at Stanford Law School, and the University of Melbourne and the Hebrew University of Jerusalem Joint Research Workshop 'The "Soft" and "Fuzzy" in Public and International Law'.

1 Vienna Convention on the Law of Treaties (VCLT) 1969, 1155 UNTS 331.

2 See, e.g., Mortenson, 'The Travaux of Travaux: Is the Vienna Convention Hostile to Drafting History?', 107 *American Journal of International Law* (*AJIL*) (2013) 780; Gardiner, 'The Vienna Convention Rules on Treaty Interpretation', in D.B. Hollis (ed.), *The Oxford Guide to Treaties* (2012) 475; Villiger, 'The Rules on Interpretation: Misgivings, Misunderstandings, Miscarriage?', in E. Cannizzaro (ed.), *The Law of Treaties beyond the Vienna Convention* (2011) 105; Sbolci, 'Supplementary Means of Interpretation', in Cannizzaro, *ibid.*, 145; Merkouris, '"Third Party" Considerations and "Corrective Interpretation" in the Interpretative Use of Travaux Prépatatoires: Is It Fahrenheit 451 for Preparatory Work?', in M. Fitzmaurice, O. Elias and P. Merkouris (eds), *Treaty Interpretation and the Vienna Convention on the Law of Treaties: 30 Years on* (2010) 75; Arsanjani and Reisman, 'Interpreting Treaties for the Benefit of Third Parties: The "Salvors' Doctrine" and the Use of Legislative History in Investment Treaties', 104 *AJIL* (2010) 597; R.K. Gardiner, *Treaty Interpretation* (2008), at 306–310; A. Orakhelashvili, *The Interpretation of Acts and Rules in Public International Law* (2008), at 382–392; Linderfalk, 'Is the Hierarchical Structure of Articles 31 and 32 of the Vienna Convention Real or Not? Interpreting the Rules of Interpretation', 54 *Netherlands International Law Review* (*NILR*) (2007) 133; Klabbers, 'International Legal Histories: The Declining Importance of Travaux Préparatoires in Treaty Interpretation?', 50 *NILR* (2003) 267; Schwebel, 'May Preparatory Work Be Used to Correct Rather Than Confirm the "Clear" Meaning of a Treaty Provision?', in Jerzy Makarczyk (ed.), *Theory of International Law at the Threshold of the 21st Century: Essays in Honor of Krzysztof Skubiscewski* (1996) 541; *Maritime Delimitation and Territorial Questions (Qatar v. Bahrain)*, Judgment, 15 February 1995, ICJ Reports (1995) 1, Dissenting Opinion of Judge Schwebel.

issue between the two positions can be seen most clearly in their respective answers to the following question: How should a treaty be interpreted in a case where the application of the general rule by itself leads to a clear and reasonable meaning of the text, while the supplementary means of interpretation support a different meaning? In these circumstances, proponents of the traditional approach will not allow recourse to supplementary means to determine the meaning of the text, whereas proponents of the corrective approach will.

The normative basis of the hierarchy debate is driven by two main controversies. The primary controversy relates to the reliability of the preparatory work as a tool for discovering the intentions of the parties to the treaty and the second controversy relates to the purpose of treaty interpretation. Thus, proponents of the traditional approach may suggest that the use of preparatory work is problematic due to its dubious reliability,[3] while proponents of the corrective approach emphasize its importance as an evidentiary tool.[4] Alternatively, it may be suggested that the intentions of the parties might not be the main goal of treaty interpretation, at least in some areas such as constitutive or human rights treaties, and, therefore, a limited role should be given to the preparatory work.[5] Although there is growing support for the less strict approach to the hierarchy between Articles 31 and 32, the traditional approach seems to remain rather dominant in the discourse on treaty interpretation.[6]

This article does not delve into the much debated question of the value of resorting to preparatory work but, rather, offers a new perspective on the debate over the hierarchy between the VCLT rules and its practical relevance. It has been suggested by several authors that, in practice, preparatory work is more widely used than what the traditional approach would allow and that exposure to the preparatory work could be the cause for its wide use.[7] Thus, for instance, Anthony Aust states that '[t]his is no

[3] See, e.g., Arsanjani and Reisman, *supra* note 2, at 602: '[W]hen a shifting cast of people representing vast organizations with many different internal and often conflicting interests makes an agreement and when other, equally complex organizations that did not participate in the making of the agreement subsequently adhere to it, the quest for intentions is utterly different. Can one then say that events anterior to the redaction of the text of the agreement are indicative of the normative expectations of the latecomers? … The very notion of the "subjective" views of a state involves a personification of a complex social organization to a degree that would make Hegel himself blush; even if the personified state is reduced to a key person, decision makers are always changing. In multilateral treaties, the quest for the "shared" subjectivities of the many states that are involved in any place other than the text of the agreement is a pursuit of the ignis fatuus.' Klabbers, *supra* note 2, at 280.

[4] See, e.g., Gardiner, *supra* note 2, at 303–310.

[5] See, e.g., Brölmann, 'Specialized Rules of Treaty Interpretation: International Organizations', in D.B. Hollis (ed.), *The Oxford Guide to Treaties* (2012) 507, at 511–512 (suggesting that in the interpretation of constitutive treaties a 'teleological approach' may be favoured over the common 'intentional approach'); Letsas, 'Strasbourg's Interpretive Ethic: Lessons for the International Lawyer', 21 *European Journal of International Law* (*EJIL*) (2010) 509, at 540: 'It follows from the above remarks that Strasbourg's interpretive ethic of dismissing drafters' specific intentions, of steering away from dictionary definitions and forays into linguistic analysis, and of applying a moral reading into the ECHR rights is fully justified by the object and purpose of human rights treaties.'

[6] See, e.g., Arsanjani and Reisman, *supra* note 2, at 600.

[7] See Mortenson, *supra* note 2, at 783. In addition to the behavioural explanation we provide here, there are several other possible explanations for a wider use of preparatory work. The less strict

doubt how things work in practice; for example, the parties to a dispute will always refer the tribunal to the *travaux*, and the tribunal will inevitably consider them along with all the other material put before it'.[8]

Aust's statement seems to imply that once interpreters are exposed to preparatory work, it is impossible for them to ignore its influence, regardless of the rules of treaty interpretation. The difficulty of ignoring information when its use is forbidden has been widely recognized in the general psychology literature[9] and in empirical legal research, but it has remained outside the research on legal interpretation.[10] Our study

---

approach to the hierarchy question provides one possible answer to the puzzle regarding the use of preparatory work in international courts. The jurisprudence may in fact endorse the alternative, less strict, approach to the use of preparatory work, even if the decisions are not explicitly based on this position. Another explanation can be found in a hermeneutic position that suggests a strong indeterminacy, allowing recourse to preparatory work to determine the meaning of a text in many cases even under the traditional view. See Merkouris, *supra* note 2, at 94. A third possibility is to attribute the wide use to the role of politics in the process of interpretation in international law. On the political dimension of interpretation in international law, see, e.g., Klabbers, 'On Rationalism in Politics: Interpretation of Treaties and the World Trade Organization', 74 *Nordic Journal of International Law* (2005) 405, at 406; Pauwelyn, 'Treaty Interpretation or Activism? Comment on the AB Report on United States: Ads and CVDs on Certain Products from China', 12 *World Trade Review* (2013) 235, at 240–241. We use politics here in a broad sense to describe different explanatory models, which emphasize the roll of individual or institutional interests in courts' decision-making processes, such as the attitudinal model, the rational choice model and even the neo-institutionalist model, which takes into account the 'law' as one of the variables of the model. For a general description of the different models, see Weinshall-Margel, 'Attitudinal and Neo-Institutional Models of Supreme Court Decision Making: An Empirical and Comparative Perspective from Israel', 8 *Journal of Empirical Legal Studies* (*JELS*) (2011) 556, at 557–559; Gillman and Clayton, 'Introduction – Beyond Judicial Attitudes: Institutional Approaches to Supreme Court Decision-Making', in C.W. Clayton and H. Gillman (eds), *Supreme Court Decision Making: New Institutionalist Approaches* (1999) 1, at 1–7. Under this line of possible explanations, a more lenient resort to the preparatory work can help courts widen their interpretive discretion while maintaining their legitimacy using reasoning that draws on the traditional view. See Pauwelyn and Elsig, 'The Politics of Treaty Interpretation: Variations and Explanations across International Tribunals', in J. Dunoff and M. Polack (eds), *International Law and International Relations: Taking Stock* (2013) 445, at 449.

8 See, e.g., A. Aust, *Modern Treaty Law and Practice* (3rd edn, 2013), at 218. See also Klabbers, *supra* note 2, at 281; Criddle, 'The Vienna Convention on the Law of Treaties in U.S. Treaty Interpretation', 44 *Virginia Journal of International Law* (*VJIL*) (2004) 431, at 440: '[A]s long as litigants bring travaux to courts' attention – as they always do – courts cannot prevent Article 31 analysis from becoming prematurely 'contaminated' by these supplementary materials.' Vandevelde, 'Treaty Interpretation from a Negotiator Perspective', 21 *Vanderbilt Journal of Transnational Law* (1988) 281, at 296–297.

9 See, e.g., Schul and Mayo, 'Discounting Information: When False Information is Preserved and When It is Not', in D.N. Rapp and J.L.G. Braasch (eds), *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences* (2014) 203; Schul, 'When Warning Succeeds: The Effect of Warning on Success in Ignoring Invalid Information', 29 *Journal of Experimental Social Psychology* (1993) 42; Wyer and Budesheim, 'Person Memory and Judgments: The Impact of Information That One is Told to Disregard', 53 *Journal of Personality and Social Psychology* (*JPSP*) (1987) 14; Anderson *et al.*, 'Perseverance of Social Theories: The Role of Explanation in the Persistence of Discredited Information', 39 *JPSP* (1980) 1037.

10 On the importance and potential of applying behavioral insights in international law research, see, e.g., Broude, 'Behavioral International Law', 163 *University of Pennsylvania Law Review* (*UPLR*) (2015) 1099; Van Aaken, 'Behavioral International Law and Economics', 55 *Harvard International Law Journal* (2014)

provides a first experimental examination of the interpretation of treaties under the VCLT rules. We have examined the ability of students and experts of international law to ignore preparatory work when it is forbidden to use such material. We have hypothesized, based on the behavioural literature, that exposure to preparatory work will affect the interpretation of a text, even under a rule that prohibits its use. We have examined this question empirically in a series of experimental studies on international law students and experts.

The results of our research suggest that exposure to preparatory work affects students' interpretation of treaties even when they are forbidden to use it, whereas no such effect was found when international law experts think that they are forbidden to use preparatory work. Our findings provide significant insights into the ongoing discussion on the relevance of legal expertise and reaffirm the relevance and efficiency of the VCLT rules. In addition, our data demonstrate that the vast majority of experts tend to support the traditional approach to treaty interpretation. This support is consistent regardless of gender, geographic location and whether the expert is a practitioner or an academic, with the exception that international human rights law (IHRL) experts tend to lend more support to the corrective approach to treaty interpretation than other experts. The implications of our study go beyond the debate over the hierarchy between the rules of treaty interpretation and can be useful for any similar debate over legal interpretation.

The article proceeds as follows. Part 2 presents the literature on the ability to ignore information in the legal context. Part 3 describes our experimental research. Part 4 discusses the results of the experiments, their possible normative implications and the limitations of the study. In Part 5, we conclude and address possible wider applications of the study.

2 Theoretical Background on the Difficulty of Deliberately Ignoring Information in the Legal Context

In the behavioural law and economics literature, the study of the ability to ignore information has focused on the ability of jurors and judges to ignore inadmissible evidence after exposure thereto. These studies have focused on civil and criminal cases, mainly examining whether judicial instructions to jurors can be an effective remedy in cases where jurors are exposed to inadmissible evidence. Most studies have supported the basic hypothesis according to which inadmissible evidence affects jurors even when they are instructed to ignore it, although there is no perfect consistency among the different studies. These studies deal mainly with criminal cases where the inadmissible evidence supports the prosecution's case, but some studies were also conducted on civil cases and with respect to inadmissible evidence that supports the defendant. In a meta-analysis of 48 studies consisting of 175 tests, Nancy Steblay

421; Chilton and Tingley, 'Why the Study of International Law Needs Experiments', 52 *Columbia Journal of Transnational Law* (*CJTL*) (2013) 173.

and colleagues found support, *inter alia*, for the hypothesis that judicial instructions do not fully eliminate the impact of the exposure to inadmissible evidence.[11] They also found that there is variability in the impact of the inadmissible evidence, depending on the type of case and the strength of the inadmissible evidence. Recent writing on this subject has accepted the major conclusions of the meta-analysis, while also considering several studies that were conducted in the years following the meta-analysis.[12] One of these studies involved an observational study on real juries' decisions, which adds to the copious literature involving experiments on mock jurors.[13] In this study, Theodore Eisenberg and Valerie Hans found a significant correlation between jurors' knowledge of prior criminal record and conviction rates in cases with relatively weak evidence, while, in cases with strong evidence, they did not find any such correlation.

While all the above-mentioned studies examined decisions made by lay persons, treaty interpretation involves interpreters who are familiar with the law in general and with international law in particular.[14] Only a few studies have been conducted on the effect of inadmissible evidence on judges. Stephen Landsman and Richard Rakos found that judges were influenced by inadmissible evidence in a tort case and that the effect of inadmissible evidence on judges was similar to its effect on mock jurors.[15] In a larger study, Andrew Wistrich, Chris Guthrie and Jeffrey Rachlinski tested judges' ability to ignore seven types of inadmissible evidence in criminal and civil cases.[16] Their results suggest that for judges, as for juries, it is difficult to ignore inadmissible evidence. Nonetheless, in two out of the seven scenarios in the article, they did not find a significant effect.[17] One of these two scenarios was a probable cause case, and, in a subsequent study that was conducted on 900 judges, they found further evidence for

[11] Steblay *et al.*, 'The Impact on Juror Verdicts of Judicial Instruction to Disregard Inadmissible Evidence: A Meta-Analysis', 30 *Law and Human Behavior* (2006) 469.

[12] Teichman and Zamir, 'Judicial Decision-Making: A Behavioral Perspective', in E. Zamir and D. Teichman (eds), *Oxford Handbook of Behavioral Economics and the Law* (2014) 664, at 671–673; Sklansky, 'Evidentiary Instructions and the Jury as Other', 65 *Stanford Law Review* (2013) 407, at 438–439 (discussing the limitations of empirical research of the issue and the exaggerated way in which this research is usually presented, but accepting to a large extent the conclusions of the meta-analysis).

[13] Eisenberg and Hans, 'Taking the Stand on Taking the Stand: The Effect of a Prior Criminal Record on the Decision to Testify and on Trial Outcomes', 94 *Cornell Law Review* (2009) 1353.

[14] For discussion of the relevant interpretation community, see Gardiner, *supra* note 2, at 109–133 (including in the relevant community of interpreters international courts and tribunals, international organizations and national legal systems).

[15] Landsman and Rakos, 'A Preliminary Inquiry into the Effects of Potentially Biasing Information on Judges and Jurors in Civil Litigation', 12 *Behavioral Science and the Law* (1994) 113.

[16] Wistrich, Guthrie and Rachlisnki, 'Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding', 153 *UPLR* (2005) 1251.

[17] The scenarios included settlement demands made during a pretrial conference; information protected by the attorney-client privilege; inadmissible sexual history in a criminal case; a presumptively inadmissible criminal record in a civil case; information obtained by the prosecution from a criminal defendant that the government had agreed not to use at sentencing under a 'cooperation agreement'; the outcome of a search involving a probable cause determination; and a criminal confession obtained during an interrogation conducted after the defendant had invoked his right to counsel.

judges' resistance to the effect of inadmissible evidence in probable cause cases.[18] To conclude, it is well established that the difficulty of ignoring irrelevant information exists in the legal context and can also affect legal experts.

Our study builds on this literature and examines whether exposure to preparatory work influences the interpretation of treaties even when the rules of interpretation forbid recourse to such materials. This study differs from previous studies in several aspects. First, research on the effect of exposure to information on decision making has focused on decisions that are not unique to legal experts, such as fact-finding and verdict rendering.[19] By contrast, our study focuses on a task that is unique to persons who possess legal expertise: legal interpretation. It has long been suggested that research should be conducted on the unique tasks of legal experts.[20] For example, Frederick Schauer has pointed specifically to legal interpretation as a potential area in which legal experts might operate differently than lay persons.[21] Second, previous studies examined cases where a binary rule applies, which either excludes the evidence in full or allows its use. In this study, even under the traditional approach, the exposure of the interpreter to the preparatory work is always allowed, while its use is limited. This exposure might weaken the perception of the recourse to preparatory work as illegitimate in comparison to cases of clear prohibition of any use of the evidence and, thus, might strengthen the effect of the exposure. Third, previous studies have examined the ability of jurors and judges to ignore information almost exclusively in criminal and civil cases, whereas our scenarios involve a domestic administrative case, an international human rights law case and a bilateral investment treaty case. Different types of cases were treated in the literature as a possible moderator variable and a useful avenue for future research.[22] Nonetheless, we do not find strong reasons to believe that the mere difference in the types of cases will make a difference.

In addition to its differences from previous studies, our study engages with the specific debate over the effectiveness of the VCLT rules. Some scholars have suggested that interpretation is more of an art than a science – namely, that the interpretive decision involves much discretion and cannot be directly attributed to the substance of the rules of interpretation.[23] These scholars point to the ambiguity and complexity

[18] Rachlinski, Guthrie and Wistrich, 'Probable Cause, Probability, and Hindsight', 8 *JELS* (2011) 72.

[19] Schauer, 'Is There a Psychology of Judging?', in David E. Klein and Gregory Mitchel (eds), *The Psychology of Judicial Decision Making* (2010) 103, at 104–105 (describing fact-finding as a task that is not unique to judges).

[20] See Teichman and Zamir, *supra* note 12, at 692–693; Schauer, *supra* note 19, at 104–105; Kahan *et al.*, '"Ideology" or "Situation Sense"? An Experimental Investigation of Motivated Reasoning and Professional Judgment', 64 *UPLR* (2016) 349.

[21] Schauer, *supra* note 20, at 104.

[22] See Steblay *et al.*, *supra* note 11, at 476, 489 (referring to the potential difference between criminal and civil cases).

[23] See, e.g., Klabbers, *supra* note 7; D. Hollis, *Art and Auto-Interpretation of Treaties, Opinio Juris* (2009), available at http://opiniojuris.org/2009/03/03/art-and-auto-interpretation-of-treaties/, but see also Linderfalk, 'Is Treaty Interpretation an Art or a Science? International Law and Rational Decision Making', 26 *EJIL* (2015) 169.

of the VCLT rules as the reason for their deficiency.[24] Usually, even those who endorse the art paradigm seem willing to accept that the VCLT rules have at least some effect on interpretation.[25] Nevertheless, there are some who still voice strong doubt as to the relevance of the VCLT rules.[26]

3 The Studies

Our research examines the effect of the rules of treaty interpretation on the interpretive process among students and experts. Experiments 1a and 1b examine the influence that exposure to preparatory work has on students enrolled in international law courses, and Experiment 2 examines its influence on international law experts. We tested three hypotheses: (i) exposure to preparatory work will affect the decisions of participants who report that they are allowed to use preparatory work in their decision; (ii) exposure to preparatory work will affect the decisions of participants who report that they are not allowed to use it; and (iii) the participants' views regarding their ability to use preparatory work under the VCLT rules of interpretation can moderate the effect of the exposure to preparatory work, even if they do not fully eliminate the effect.

A *Experiment 1a: The Effect of Exposure to Preparatory Work on International Law Students*

1 Participants

We recruited 269 law students in their advanced years of study (mean age 24.8, standard deviation (SD) = 5, 151 females) who were enrolled in international law courses at the Hebrew University of Jerusalem, the Herzliya Interdisciplinary Center, the Rishon LeZion College of Management and Sha'arei Mishpat College. All participants in our study volunteered to answer our questionnaire without any compensation.

2 Procedure

We distributed our questionnaires to students in the early stages of their international law courses. We asked them to read and respond to each of the questions and to do so independently. The rooms were silent during the administration of the questionnaires. We did not ask the participants for any identifying details and informed them that participation was entirely voluntary. Each student who agreed to participate received a questionnaire in Hebrew that included the relevant VCLT rules and two scenarios.[27] The scenarios are based on two real cases, *Witold Litwa v. Poland*, which was held

[24] Pauwelyn and Elsig, *supra* note 7, at 448; Klabbers, *supra* note 7, at 407–408; See M. Koskenniemi, *From Apology to Utopia* (rev. edn, 2005), at 333–345.

[25] Klabbers, *supra* note 7, at 411.

[26] See Koskenniemi, *supra* note 24, at 333–345.

[27] The version of the rules that the participants were exposed to is a redacted version of the VCLT rules that includes only Arts 31(1), 31(2) and 32. We decided to use this redacted version of the rules, which maintains the core of the hierarchy question and renders the VCLT rules less complicated.

before the European Court of Human Rights (ECtHR),[28] and *Afu et al. v. Commander of IDF Forces*, which was held before the Israeli High Court of Justice.[29] In both cases, the court was divided in relation to the proper use of the VCLT rules, and both cases have been discussed in the international law literature with regard to their use of the preparatory work of the relevant conventions.[30] These two cases demonstrate that the hierarchy question is not merely a theoretical issue but can also have practical influence on real cases.

Each participant was randomly assigned to one of four experimental conditions. The four versions of the questionnaire varied in terms of the exposure to relevant preparatory work (each participant was exposed to preparatory work in one of two cases and was not exposed to preparatory work in the other case) and in terms of the order in which the two cases appeared.[31] After reading each case, participants were asked to decide whether the relevant article was violated and to justify their decision. Following the participants' decisions in both cases, they were asked to answer several questions regarding treaty interpretation, such as whether a hierarchy exists between the general rule and the supplementary means of interpretation and whether recourse to preparatory work is allowed under the circumstances of the specific case. Lastly, participants were asked whether they had previously been exposed to the cases and to provide demographic information.

[28] ECtHR, *Litwa v. Poland,* Appl. no. 26629/95, Judgment of 4 April 2000. Decision available online at http://hudoc.echr.coe.int/.

[29] HCJ 785/87, *Afu v. Commander of IDF Forces*, PD 42(2) 4 (Israel 1988).

[30] For discussion of the use of the rules of interpretation in the *Litwa* case, see Gardiner, *supra* note 2, at 501: '[I]n reaching this interpretation, the Court ostensibly relied on context as well as the treaty's object and propose to displace an apparently unequivocal ordinary meaning. It then "confirmed" this view by reference to the provision's preparatory work, noting that the commentary on the preliminary draft acknowledged the right of States to take measures to combat vagrancy and drunkenness. Although this application of the Vienna rules seems fully in keeping with their proper use, it is difficult not to conclude that the consideration of the preparatory work before formal application of the general rule convinced the Court of the correct interpretation. Further, it seems inevitable that courts and tribunals commonly consider preparatory work before formulating their judgment or award. Only in the loosest sense is this process "confirming" a meaning established by the general rule, even if (as in the example above) care is taken to construct the interpretation giving respect to the structure of the Vienna rules'). For a discussion of the use of the rules of interpretation in the *Afu* case, see, e.g., Margalit and Hibbin, 'Unlawful Presence of Protected Persons in Occupied Territory? An Analysis of Israel's Permit Regime and Expulsions from the West Bank under the Law of Occupation', 13 *Yearbook of International Humanitarian Law* (2010) 245, at 255–260; Y. Dinstein, *The International Law of Belligerent Occupation* (2009), at 165; D. Kretzmer, *The Occupation of Justice: The Supreme Court of Israel and the Occupied Territories* (2002), at 148–152; T. Meron, *Human Rights and Humanitarian Norms as Customary Law* (1989), at 48–49, n. 131: 'The Language of Article 49(1) is clear and categorical. It prohibits both individual and collective deportations, regardless of their motive. Recourse to supplementary means of interpretation (Article 32 of the Vienna Convention) was, therefore, not justified.' Though based on the real cases, the experimental scenarios were not identical to these cases, neither in their factual description nor in the alleged preparatory work that was provided to the participants. The changes were intended to make the intended contradiction between the relevant article textual interpretation and the preparatory work more striking than the actual cases. Therefore, it is not possible to directly infer anything from the results of the current study with regard to the application of the VCLT rules of interpretation in the real cases.

[31] We checked for order effects and did not find any significant effect of order in any of our experiments.

3 Scenario 1: Detention of a Drunk Man (Litwa)

(a) The scenario

The first scenario was loosely based on the *Litwa* case before the ECtHR. The participants were asked to imagine that they were sitting as judges at the ECtHR. They were given a description of a case involving a man who drank alcohol for the first time in his life. He then entered a post office in Krakow and insulted the post office clerks. Policemen who arrived at the scene took him into custody and detained him for 6.5 hours until he sobered up and was released. The participants were then informed that the European Convention on Human Rights (ECHR) allows the detention of persons only under a closed list of grounds for detention. They were presented with Article 5(1)(e) of the ECHR, which allows detention 'for the prevention of the spreading of infectious diseases, of persons of unsound mind, alcoholics or drug addicts or vagrants'.[32] Lastly, they were informed that the Polish government had argued that the detention lies within the circumstances described in Article 5(1)(e), while the attorneys of the applicant had argued that it applies to alcoholics, and since the applicant was only temporarily drunk and not an alcoholic, his arrest was unlawful. The experimental group received, in addition to the scenario, part of the preparatory work of the convention, which states that the article was designed to protect public order and that the text of the article covers the right of the state parties to the convention to take necessary measures to protect public order from interferences arising from public drunkenness. The control group was not exposed to the preparatory work of the convention. After reading the scenario and, in the experimental condition, the preparatory work, all participants were asked to determine whether the rights of the applicant under the ECHR had been violated and to justify their decision.

(b) Results

We excluded from the analysis two participants who reported that they had been exposed to the *Litwa* case prior to the experiment and seven participants who did not answer the question about previous exposure to the case.

Hypothesis 1

First, we tested the influence of exposure to preparatory work on the decisions of participants who reported that they were allowed to use preparatory work.[33] The results are presented in Table 1. Among the participants who were exposed to the preparatory work, 19.3 per cent determined that Article 5 of the ECHR had been violated,

[32] Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR) 1950, 213 UNTS 222.

[33] These results include participants who held to the corrective approach to interpretation and those who held to the traditional approach but determined that in the specific scenario it is possible to use the preparatory work to determine the meaning under the conditions of the articles of the VCLT.

Table 1: Decisions of Participants Who Reported That Resorting to Preparatory Work Was Allowed under This Scenario

	No violation (%)	**Violation (%)**
Exposed to preparatory work	88 (80.7)	21 (19.3)
Not exposed to preparatory work	56 (54.9)	46 (45.1)
Total	144 (68.2)	67 (31.8)

compared to 45.1 per cent of the participants who were not exposed to the preparatory work. The difference in proportions is significant (X^2 [1, N = 211[34]] = 16.225, $p < 0.001$). As we hypothesized, participants who were exposed to preparatory work were less likely to determine that Article 5 of the ECHR was violated, compared to participants who were not exposed to the preparatory work.[35]

Hypothesis 2

In order to test the effect of exposure to preparatory work on participants who believed that resorting to preparatory work was not allowed under the circumstances of the scenario, we analysed only the results of participants who held to the traditional approach[36] and determined that resorting to preparatory work was not allowed under this scenario. The results are presented in Table 2. Among these participants, 43.5 per cent of those who were exposed to preparatory work determined that Article 5 of the ECHR had been violated, whereas 91.3 per cent of those who were not exposed to preparatory work determined that such a violation had occurred. The difference in proportions is significant ($p = 0.001$, two-tailed Fisher's exact test). Again, as we hypothesized, participants who had been exposed to preparatory work were more likely to determine that there had been no violation of Article 5 of the ECHR, compared to participants who were not exposed to preparatory work.[37]

A logistic regression did not yield any statistically significant effect of political opinion, age or gender on the decision whether Article 5 of the ECHR was violated.[38] We did not

[34] Three participants did not answer whether or not there had been a violation.

[35] Inclusion of the participants who had been exposed to the case did not change the pattern of results and they remained significant (X^2 [1, N = 219] = 15.938, $p < 0.001$).

[36] That is, participants who reported that in cases of a clear and reasonable interpretive result following the application of Art. 31, it was impermissible to use preparatory work to determine the meaning of the text.

[37] Inclusion of the participants who had been exposed to the case did not change the pattern of results, and they remained highly significant.

[38] We also examined a possible indirect effect of exposure to preparatory work. We hypothesized that in addition to its direct effect on the decision regarding the violation of the relevant treaty articles, the exposure to the preparatory work can affect the determination of those participants who hold to the traditional approach whether the meaning of the relevant article is ambiguous. See Vandevelde, *supra* note 8, at 297 (suggesting an effect of the exposure on determinations of ambiguity). We found no such effect in the students' study detention scenario and in the two experts' scenarios. We did find an effect in the deportation scenario ($p = 0.005$, two-tailed Fisher's exact test). We have no simple explanation for the mixed results, and further research should be conducted to examine a possible indirect effect to the exposure to information in legal interpretation.

Table 2: Decisions of Participants Who Reported That Resorting to Preparatory Work Was Not Allowed under This Scenario

	No violation (%)	**Violation (%)**
Exposed to preparatory work	13 (56.5)	10 (43.5)
Not exposed to preparatory work	2 (8.7)	21 (91.3)
Total	15 (32.6)	31 (67.4)

find these or other variables, such as years of post-graduate legal experience and field of expertise, to have a significant effect in any of the other experiments of our study.

Hypothesis 3

In order to test the third hypothesis, according to which the participants' views regarding their ability to use preparatory work under the VCLT rules of interpretation can moderate the effect of the exposure to preparatory work on the decision on whether Article 5 of the ECHR was violated, we compared the decisions of the participants who determined that resorting to the preparatory work was allowed under this scenario and those of the participants who determined that it was prohibited. In line with our hypothesis, when exposed to the preparatory work, participants who reported that the use of preparatory work was allowed determined that the article was not violated more than those participants who determined that it was prohibited. Nevertheless, a logistic regression on participants' decisions did not reveal a significant interaction between exposure to preparatory work and their position on resorting to preparatory work under the specific scenario.

4 Scenario 2: Deportation (Afu)

(a) The scenario

Scenario 2 is loosely based on the *Afu* case before the Israeli Supreme Court. The participants were asked to imagine that they were sitting as judges in a case where the Israeli government wishes to deport outside of the Occupied Territories a central member of Hamas, who was found to be involved in planning terror attacks against civilians. His attorneys argue that his deportation will violate Article 49(1) of the Geneva Convention IV, which states that '[i]ndividual or mass forcible transfers, as well as deportations of protected persons from occupied territory to the territory of the Occupying Power or to that of any other country, occupied or not, are prohibited, regardless of their motive'.[39] The experimental group (which was the control group in the *Litwa* scenario) received in addition to the scenario a text of the preparatory work, which suggested that the article was drafted in light of the mass deportations conducted by the Nazis during World War II and that the article was designed to prevent such horrible instances in the future. Next, participants were asked to determine whether deportation of the Hamas member violates Article 49(1) of Geneva Convention IV.

[39] Geneva Convention IV Relative to the Protection of Civilian Persons in Time of War 1949, 75 UNTS 287.

(b) Results

We excluded from the analysis 81 participants who reported that they had been exposed to the case prior to the experiment and seven participants who did not answer the question about previous exposure.[40]

Hypothesis 1

First, we tested the influence of exposure to preparatory work on the decisions of participants who reported that they were allowed to use preparatory work. The results are presented in Table 3. In total, 37.7 per cent of the participants who were exposed to the preparatory work determined that Article 49(1) of Geneva Convention IV had been violated, compared to 66.7 per cent of the participants who were not exposed to the preparatory work. The difference in proportions is significant (X^2 [1, N = 133[41]] = 11.13, $p = 0.001$). As we hypothesized, participants who were exposed to the preparatory work were more likely to determine that there had been no violation of Article 49(1) of Geneva Convention IV, compared to participants who were not exposed to the preparatory work.[42]

Table 3: Decisions of Participants Who Reported That Resorting to the Preparatory Work Was Allowed under This Scenario

	No violation (%)	**Violation (%)**
Exposed to preparatory work	38 (62.3)	23 (37.7)
Not exposed to preparatory work	24 (33.3)	48 (66.7)
Total	62 (46.6)	71 (53.4)

Hypothesis 2

As in Scenario 1, we tested the effect of exposure to preparatory work on participants who believed that resorting to the preparatory work was not allowed under the circumstances of the scenario. The results are presented in Table 4. Among the participants who held to the traditional approach and determined that resorting to preparatory work was not allowed under this specific scenario, 72.7 per cent of those who were exposed to the preparatory work determined that Article 49(1) of Geneva Convention IV had been violated, compared to 95.7 per cent of the participants who were not exposed to the preparatory work. The difference in proportions is significant ($p = 0.047$, two-tailed Fisher's exact test). Again, as we hypothesized, participants who were exposed to preparatory work were more likely to determine that there was

[40] *Afu* is a well-known case in Israel and is often taught in international law courses. This may explain the relatively high number of participants who had been exposed to the case.

[41] Three participants did not answer whether or not there had been a violation.

[42] Inclusion of the participants who had been exposed to the case did not change the pattern of results, and they remained significant.

Table 4: Decisions of Participants Who Reported That Resorting to Preparatory Work Was Not Allowed under This Scenario

	No violation (%)	**Violation (%)**
Exposed to preparatory work	6 (27.3)	16 (72.7)
Not exposed to preparatory work	1 (4.3)	22 (95.7)
Total	7 (15.6)	38 (84.4)

no violation of Article 49(1) of Geneva Convention IV, compared to participants who were not exposed to preparatory work.[43]

Hypothesis 3

As in the *Litwa* scenario, we examined whether there was a difference between the two groups of participants, namely, those who determined that they were allowed to resort to the preparatory work in order to determine the meaning of the text in this scenario and those who determined that they were prohibited from using the preparatory work. Participants' decisions on whether Article 5 of the ECHR was violated were in line with our hypothesis – namely, when exposed to the preparatory work, participants who reported that the use of preparatory work was allowed determined that the article was violated more than did those participants who determined that it was prohibited. Nevertheless, a logistic regression on participants' decisions did not reveal a significant interaction between exposure to preparatory work and their position on resorting to preparatory work under the specific scenario. Since the results in both cases went in the direction of Hypothesis 3 but did not reach statistical significance, we conducted a second experiment to test whether the rules of interpretation could moderate the effect of exposure to preparatory work.

B *Experiment 1b: Can the Rules of Interpretation Moderate the Effect of Exposure to Preparatory Work?*

The results of our first experiment demonstrate that exposure to preparatory work may affect treaty interpretation even under a legal regime that prohibits its use. We did not find that the participants' views on their ability to use preparatory work under the VCLT rules had a moderating effect, although participants who reported that they were allowed to use the preparatory work decided more frequently than the other group that the relevant articles under both scenarios had been violated. Our second study directly tested the ability of interpretive rules to moderate the effect of exposure to preparatory work.

[43] Inclusion of the participants who had been exposed to the case did not change the pattern of results, and they remained significant ($p = 0.02$).

1 Participants

We recruited a total of 208 students in their advanced years of study (mean age 26.75, SD = 8.7, 114 females[44]) in international law courses at the Hebrew University of Jerusalem, Sha'arei Mishpat College and Tel Aviv University, who did not participate in Experiment 1a.

2 Procedure

The procedure was similar to that of Experiment 1a, with two changes. First, in order to test the influence of the interpretation rules on participants' interpretation dynamics, we presented the participants with invented rules of interpretation before they read the scenarios. Half of the participants received a version of the rules that prohibited the use of preparatory work in all circumstances, whereas the other half received a version of the rules that allowed the use of preparatory work in all circumstances. Second, all of the participants were exposed to the preparatory work in both scenarios. It is important to note that this experiment uses a binary rule of interpretation, which either fully prohibits or fully allows the use of the rules. This leaves open the possible effect of more nuanced rules, such as Articles 31 and 32 of the VCLT.[45]

3 Results of Scenario 1: Detention of a Drunk Man

We excluded from the data analysis 56 participants who reported that they were previously exposed to the VCLT rules. We tested whether the different interpretation rules influenced participants' decisions on whether Litwa's detention was lawful. The results are presented in Table 5. Among the participants who received the rules that allowed the use of preparatory work without limitations, 21.2 per cent determined that Article 5(1)(e) of the ECHR had been violated, compared to 43.2 per cent of the participants who received the rules that prohibited the use of preparatory work in all circumstances. The difference in proportions is significant (X^2 [1, N = 147[46]] = 7.919, $p < 0.005$). As we hypothesized, participants who received the rules that allowed the use of preparatory work were more likely to determine that there had been no violation of Article 5(1)(e) of the ECHR, compared to participants who received the rules that prohibited the use of preparatory work.[47]

Table 5: Participants' Decisions on Whether Article 5(1)(E) Was Violated

	No violation (%)	**Violation (%)**
Rules allow use of preparatory work	52 (78.8)	14 (21.2)
Rules prohibit use of preparatory work	46 (56.8)	35 (43.2)
Total	98 (66.7)	49 (33.3)

[44] Seven participants did not report their gender.

[45] See discussion in part 2 of this article.

[46] Five participants did not answer whether or not there had been a violation.

[47] Inclusion of the participants who had been exposed to the rules did not change the pattern of results, and they remained significant.

4 Results of Scenario 2: Deportation

In addition to the exclusion of participants who were exposed to the VCLT, we excluded from the analysis 12 participants who reported that they had been exposed to the *Afu* case. We tested whether the different interpretation rules influenced participants' decisions on the legality of the deportation. The results are presented in Table 6. Among the participants who received the rules that allowed the use of preparatory work without limitations, 42.9 per cent determined that Article 49(1) of Geneva Convention IV had been violated, compared to 75 per cent of the participants who received rules that prohibited the use of preparatory work. The difference in proportions is significant (X^2 [1, N = 135[48]] = 14.464, $p < 0.001$). As we hypothesized, participants who received the rules that allowed the use of preparatory work were more likely to determine that there had been no violation of Article 49(1) of Geneva Convention IV compared to participants who received the rules that prohibited the use of preparatory work.[49]

Table 6: Participants' Decisions on Whether Article 49(1) Was Violated

	No violation (%)	**Violation (%)**
Rules allow to use preparatory work	36 (57.1)	27 (42.9)
Rules prohibit to use preparatory work	18 (25.0)	54 (75.0)
Total	54 (40.0)	81 (60.0)

C *Experiment 2: International Law Experts*

The two experiments in our study of law students indicate that students do not completely ignore information they should not use while making legal decisions. Legal experts might act differently. The interpretation of legal rules is one of the unique tasks of legal experts. In the context of international law, in which there is no institutional interpretive hierarchy, interpretation is central to the work of a variety of legal actors.[50] Although conducting experiments on law students as proxies for legal experts is an accepted practice in empirical legal scholarship,[51] we conducted an additional study on international law experts in order to examine whether the effect of the exposure replicates on legal experts. We examined two main hypotheses regarding the influence that exposure to preparatory work has on experts' decisions: (i) exposure to preparatory work affects decisions of experts who report that they are allowed to use preparatory work in the relevant case and (ii) exposure to preparatory work affects decisions of experts who report that they are not allowed to use preparatory work.

48 Five participants did not answer whether or not there had been a violation.

49 Inclusion of the participants who had been exposed to the case did not change the pattern of results, and they remained significant.

50 Gardiner, *supra* note 2, at 110: 'Thus internationally, issues over treaty interpretation will commonly be a matter for discussion, negotiation, and agreement between states or for resolution within an international organization, with judicial or arbitral determination covering only a small minority of cases.'

51 See Teichman and Zamir, *supra* note 12, at 692; Zamir and Ritov, 'Loss Aversion, Omission Bias, and the Burden of Proof in Civil Litigation', 41 *Journal of Legal Studies* (2012) 165, at 183–184.

1 Participants

For the experts study, we recruited 212 international law practitioners and academics[52] (mean age 44.8, SD = 12.6, 66 females[53]). In total, 157 of them were academics, 48 were practitioners (including governmental, military and private practitioners) and four were international judges and arbitrators.[54] The experts had an average of 15.1 years of post-graduate experience (SD = 11.8) and a variety of areas of expertise (50 had experience in IHRL, 49 in international economic law, 17 in international humanitarian law, 20 in international criminal law and 74 had other areas of expertise).[55] The participants were also geographically diverse (76 participants were from Europe, 65 from the USA and Canada, 20 from Israel, 17 from Australia and New Zealand, 7 from Asia, 5 from Africa and 5 from Latin America).[56]

2 Procedure

We recruited the experts online by posting requests to participate in the study via several interest groups of the American and European societies of international law; we distributed emails to Israeli government and military lawyers; and we sent personal emails to international law academics and practitioners in top international law firms, based on lists of members of international law societies, lists of areas of expertise of faculty in law school websites and lists of partners in top international law firms.[57] Experts who agreed to participate received an online questionnaire that included two scenarios. Scenario 1 was similar to the

[52] To enable large enough sample, we included participants with any practical or academic post-graduate experience.

[53] Three participants did not report their gender.

[54] Three participants did not report their organization. In all sub-categories, in cases of more than one answer, we used only the first answer.

[55] Two participants did not report their area of expertise.

[56] Seventeen participants did not report their nationality.

[57] It is very difficult to find the population of international law experts since there is no systematic data about them. For a similar difficulty, see Grossman, 'Shattering the Glass Ceiling in International Adjudication', 56 *VJIL* (2016) 339 (discussing the difficulty to find the pool of women international law lawyers). Because we do not have a complete list of the whole population of international law experts, using probability sampling was impractical and infeasible for us. We used, as described, non-probability sampling, and we included in our analysis all of the experts who agreed to answer the questionnaire. Consequently, we have no way to estimate our sampling error. This makes it more difficult to generalize our findings to the entire population of international law experts or to the narrower population of international judges. Specifically, we recruited the experts to our study based on English contact details online, through the American and European societies international law. Thus, Asian, Latin America and African experts are potentially underrepresented in our study. Nonetheless, the sources that we used as the basis of our recruitment strategy, especially the member directory of the American Society of International Law, seem to contain the main available lists of international law experts, and the use of similar types of non-probabilistic sampling is rather common in studies on legal experts. See Kahan *et al.*, *supra* note 20, at 375–376; Zamir and Ritov, *supra* note 55, at176; Feldman, Schurr and Teichman, 'Anchoring Legal Standards', 13 *JELS* (2016) 298, at 313. In addition, as mentioned, we did not find an effect of the age, gender, political opinion, years of post-graduate experience and field of expertise on the experts' decisions.

Litwa scenario[58] in the students' study.[59] Scenario 2 involved an invented bilateral investment treaty (BIT) arbitration dispute. We chose to use this case for two reasons. First, we assumed that most experts would be familiar with the *Afu* case since it is a widely known case in the international law community. Second, we wanted to use a case from international economic law as it is an area that is usually regarded as being less political than international human rights law or international humanitarian law. As in the first study, each participant was exposed to preparatory work in one case and was not exposed to preparatory work in the other case. After reading each case, participants were asked to determine whether there had been a violation of the relevant treaty and to justify their decision. Following their decisions in both cases, participants were asked to answer several questions regarding treaty interpretation and previous exposure to the cases and to provide demographic information. The order in which the scenarios appeared was counterbalanced between participants.[60]

3 Scenario 1: Detention of a Drunk Man (Litwa)

(a) Results

We excluded from the analysis 15 participants who reported that they had been exposed to the *Litwa* case prior to the experiment.

Hypothesis 1

We tested whether experts who reported that they were allowed to use preparatory work were influenced by the exposure to the preparatory work in their decision on whether Litwa's detention was lawful. The results are presented in Table 7. In total, 16.1 per cent of the experts who were exposed to the preparatory work determined that Article 5 of the ECHR had been violated, compared to 45.9 per cent of the experts who were not exposed to the preparatory work. The difference in proportions is significant (X^2 [1, N = 117[61]] = 12.01, $p < 0.001$).[62] As we hypothesized, participants who were exposed to the preparatory work were more likely to determine that there had been no violation of Article 5 of the ECHR, compared to participants who were not exposed to the preparatory work.

[58] The scenario was translated to English.

[59] Following the difference between the decisions of experts and of law students, after collecting the answers of 100 experts, we changed a few minor details in the scenarios to make the contrast between the text and the preparatory work even more striking (e.g., we extended the detention time from 6.5 hours to 12 hours in the drunk man scenario). The pattern of the results in both versions is similar, therefore we report their combined results.

[60] We checked for order effects and found no significant effect for the order of the scenario in both scenarios.

[61] Six participants did not answer whether or not there had been a violation.

[62] Inclusion of the participants who were exposed to the *Litwa* case did not change the pattern of results (X^2 [1, N = 125] = 9.04, $p = 0.003$).

Table 7: Decisions of Experts Who Reported That Resorting to Preparatory Work Was Allowed under This Scenario

	No violation (%)	**Violation (%)**
Exposed to preparatory work	47 (83.9)	9 (16.1)
Not exposed to preparatory work	33 (54.1)	28 (45.9)
Total	80 (68.4)	37 (31.6)

Hypothesis 2

Second, we tested whether the effect of exposure to preparatory work was similar among experts who believed that resorting to preparatory work was not allowed under the circumstances of the scenario. Therefore, we analysed only the results of experts who held to the traditional approach and who, in addition, determined that resorting to the preparatory work was not allowed under this specific scenario. The results are presented in Table 8. Among these experts, 69.7 per cent of those who were exposed to the preparatory work determined that Article 5 of the ECHR had been violated, compared to 62.2 per cent of those who were not exposed to the preparatory work who determined that such a violation had occurred. The difference in proportions is not statistically significant (X^2 [1, N = 70[63]] = 0.44, p = 0.507).[64] Thus, in contrast to all students and to those experts who reported that they were allowed to use preparatory work, we did not find an effect of the exposure to preparatory work on experts who reported that they are not allowed to use such materials.

Next, we examined whether there was a difference between the decisions of the two groups of experts – those who determined that they were allowed to resort to preparatory work in order to determine the meaning of the text under this scenario and those who determined that such a use of the preparatory work was prohibited. Using a logistic regression on the experts' decisions on whether Article 5 of the ECHR was violated, we found a significant interaction between exposure to the preparatory work and the experts' position on resorting to preparatory work under the specific scenario (B = 0.456, Wald = 7.292, p = 0.007). This means that experts who reported that they were allowed to use preparatory work were influenced by its substance significantly more than were experts who reported that they were not allowed to use it.

Whereas in Experiment 1a we found that exposure to preparatory work decreased the likelihood of students determining that Article 5 of the ECHR had been violated, we did not find a significant difference between the decisions of the experts who were exposed to the preparatory work and the decisions of the experts who were not exposed to the preparatory work. Using a logistic regression, we tested the interaction between the types of participants (students versus experts) and the exposure to preparatory work (exposed versus not exposed) on the decision on whether there

[63] Four participants did not answer whether or not there had been a violation.

[64] Inclusion of the participants who were exposed to the *Litwa* case did not change the pattern of results (p = 0.603).

Table 8: Decisions of Experts Who Reported That Resorting to Preparatory Work Was Not Allowed under This Scenario

	No violation (%)	**Violation (%)**
Exposed to preparatory work	10 (30.3)	23 (69.7)
Not exposed to preparatory work	14 (37.8)	23 (62.2)
Total	24 (34.3)	46 (65.7)

had been a violation of Article 5 of the ECHR. We did not find a difference between the two types of participants who reported that resorting to preparatory work was allowed (B = –0.063, Wald = 40.002, $p = 0.643$). However, we did find a significant difference between the two types of participants who reported that resorting to preparatory work was not allowed (B = 0.738, Wald = 8.855, $p = 0.003$). The interaction between the type of respondents and exposure to preparatory work suggests that there is a qualitative difference between the interpretive processes of students and experts. There are some limitations to this comparison between students and experts that have to be taken into account. The students experiment was run in Hebrew in a classroom using pens and paper, while the experts experiment was run online in English.[65]

4 Scenario 2: BIT

(a) The scenario

Scenario 2 is an invented BIT arbitration case. The case involves a dispute between Sweden and a Slovenian firm called GGL. GGL signed a contract with the Swedish government regarding acquisition of an iron mine. The Swedish government transferred the shares of the state-owned company that owned the mine to GGL a year and a half after the payment, while the contract required an immediate transfer of the shares. GGL subsequently filed a request for arbitration to the International Centre for the Settlement of Investment Disputes (ICSID). The firm argued that the alleged breach of contract could be brought before ICSID under Articles 8(3) and 20 of the BIT between Sweden and Slovenia. Article 8(3) determines that '[e]ach

[65] Although previous studies indicate that thinking in a foreign language reduces decision biases compared to thinking in a native language (see Keysar, Hayakawa and An, 'The Foreign-Language Effect: Thinking in a Foreign Tongue Reduces Decision Biases', 23 *Psychological Science* (2012) 661), in the current study we did not find any difference between judgments of legal experts who speak English as a native language and legal experts who speak English as a foreign language. The experts experiments were run online, which has some important advantages, such as cultural diversity and a better ability to ensure participants' anonymity. Research demonstrates that, for a large part, data that have been obtained online are at least as reliable as data obtained via traditional methods. See, e.g., Germine *et al.*, 'Is the Web as Good as the Lab? Comparable Performance from Web and Lab in Cognitive/Perceptual Experiments', 19 *Psychonomic Bulletin and Review* (2012) 847. Therefore, we do not believe that the different experimental setting accounts for the different pattern of results.

Contracting Party shall observe any other obligation it has assumed with regard to investments in its territory by nationals or companies of the other Contracting Party', and Article 20 states that '[t]he investor shall be entitled to submit any dispute premised on this convention to the International Center for the Settlement of Investment Disputes (ICSID)'. Sweden has argued that the BIT was not intended to cover mere contractual disputes, and, therefore, the alleged breach cannot be brought before ICSID.

The experimental group (which was the control group in the *Litwa* scenario) received, in addition to the scenario, a text of the preparatory work according to which the BIT had a previous draft of Article 8(3), which explicitly stated that it covers 'any contractual obligations between the contracting party and nationals or companies of the other contracting party'. The article was changed to its current form due to a statement by the head of the Slovenian delegation that Slovenia 'objects to this wording since regular contractual obligations, in contrast to obligations which reflect the role of the state as a sovereign, should not be part of the convention protection'. The participants were asked to determine whether the alleged contractual breach violated Article 8(3) of the BIT. We hypothesized that the experimental group would tend to decide that the article had not been violated more than the control group would.

(b) Results

Hypothesis 1

First, we tested whether participants who reported that they were allowed to use preparatory work were influenced by the exposure to the preparatory work in their decision on whether Article 8(3) of the BIT had been violated. The results are presented in Table 9. In total, 24.6 per cent of the participants who were exposed to the preparatory work determined that Article 8(3) had been violated, compared to 67.3 per cent of the participants who were not exposed to the preparatory work. The difference in proportions is significant (X^2 [1, N = 116[66]] = 6.72, $p < 0.001$). As we hypothesized, participants who were exposed to the preparatory work were more likely to determine that there had been no violation of Article 8(3) compared to participants who were not exposed to the preparatory work.

Table 9: Decisions of Experts Who Reported That Resorting to The Preparatory Work Was Allowed under This Scenario

	No violation (%)	**Violation (%)**
Exposed to preparatory work	46 (75.4)	15 (24.6)
Not exposed to preparatory work	18 (32.7)	37 (67.3)
Total	64 (55.2)	52 (44.8)

[66] Two participants did not answer whether or not there had been a violation.

Hypothesis 2

Second, we tested whether the effect of exposure to the preparatory work was similar among participants who believed that resorting to the preparatory work was not allowed under the circumstances of the scenario. Therefore, we analysed only the results of participants who held to the traditional approach and who, in addition, determined that resorting to the preparatory work was not allowed under this specific scenario. The results are presented in Table 10. Among these participants, 85.1 per cent of those who were exposed to the preparatory work determined that Article 8(3) had been violated, and 84.6 per cent of those who were not exposed to the preparatory work determined that such a violation had occurred. The difference in proportions is not statistically significant (X^2 [1, N = 86[67]] = 0.004, p = 0.950). As in the *Litwa* scenario, we did not find an effect of the preparatory work on experts who reported that they are not allowed to use such materials.

Table 10: Decisions of Experts Who Reported That Resorting to The Preparatory Work Was Not Allowed under This Scenario

	No violation (%)	**Violation (%)**
Exposed to preparatory work	7 (14.9)	40 (85.1)
Not exposed to preparatory work	6 (15.4)	33 (84.6)
Total	13 (15.1)	73 (84.9)

In this BIT scenario, as in the *Litwa* scenario, we examined whether there was a difference between the two groups of experts – those who determined that it was allowed to resort to the preparatory work and those who determined that such use of the preparatory work was prohibited. Using a logistic regression on experts' decisions on whether Article 8(3) of the BIT had been violated, we found a significant interaction between exposure to the preparatory work and experts' position on resorting to preparatory work under the specific scenario (B = 0.470, Wald = 6.592, p = 0.010). This means, again, that experts who reported that they were allowed to use preparatory work were influenced by its substance significantly more than were experts who reported that they were not allowed to use it.

5 *Experts' Positions on the Hierarchy of the VCLT Rules*

Conducting this study on legal experts enabled us to collect descriptive data on current positions in the international law community on the hierarchy debate. We think that our descriptive data can provide an important insight into the actual positions of international law experts on the hierarchy question.[68] The results are presented in Table 11. Our data suggest that among international law experts, the traditional approach is still the dominant approach to treaty interpretation. In total, 69.3 per cent of the experts reported that they held to the traditional approach. A regression analysis did not yield

[67] Eight participants did not answer whether or not there had been a violation.

[68] Taking into account the limitations of our sample, see note 61 above.

any significant effect that gender, political opinion, geographic location, exposure to preparatory work, being an academic or practitioner and being an international economic law expert may have on the experts' positions on the hierarchy debate. We found a significant effect of the years of post-graduate experience (B = 0.046, Wald = 5.832, $p = 0.016$), such that experts with more experience tended to lend more support to the traditional approach. We also found a significant effect of IHRL expertise (B = –1.059, Wald = 7.395, $p = 0.007$). IHRL experts supported the corrective approach more than other international law experts (46 per cent versus 25.9 per cent respectively).[69]

Table 11: Experts' Positions on the Hierarchy Debate

	Traditional approach (%)	**Corrective approach (%)**
IHRL experts	27 (54.0)	23 (46.0)
Other experts	120 (74.1)	42 (25.9)
Total	147 (69.3)	65 (30.7)

4 Discussion

A *General Discussion*

This article describes the results of two studies we conducted in order to explore people's ability to ignore information in the context of treaty interpretation. One study focused on students in international law courses and one on legal experts. The results of the students study demonstrate that the VCLT rules can moderate the effect of exposure to preparatory work but that they do not eliminate it altogether. These results indicate that the general phenomenon of the difficulty of ignoring information should also be considered in the context of legal interpretation. The results of the experts study reveal a different pattern. The experts study demonstrates that international law experts can potentially resist the influence of exposure to information under a rule that prohibits its use. While experts who reported that they were allowed to use preparatory work in order to determine the meaning of the text in the specific scenario were more likely to interpret the relevant article in line with the substance of the preparatory work, no such effect was found among experts who reported that such use was not allowed.

These results can contribute to the international law literature as well as to the empirical legal research on decision making. Our descriptive data emphasize the continuing prominence of the traditional approach to treaty interpretation. The greater support for the traditional approach among experts with more experience is not surprising due to the relatively new rise of the corrective approach in the international law community. The main exception to the strong support for the traditional approach is human rights experts, who tend to support the traditional approach less than other international law experts. These findings seem to be contrary to the suggestion that human rights treaty interpretation might be different from other

[69] We reported the positions of international human rights law experts compared to all other experts. Dividing the other experts into sub-groups leads to similar results.

branches of international law by giving less weight to the intention of the parties in the process of treaty interpretation for an approach that gives primacy to teleological interpretation.[70] However, we think that both tendencies can be explained by suggesting that human rights lawyers tend to be less formalistic than their counterparts. Teleological interpretation is often considered to be a more expansive and less formalistic interpretive method than other interpretive methods[71] as does the support for the corrective approach to the VCLT rules. Both positions give the interpreter more discretion and lend support to the perception of human rights lawyers as result-driven jurists.[72]

Our descriptive data, together with the experimental results of the experts study, reaffirms the relevance of the hierarchy debate. In particular, it demonstrates that preparatory work can play a significant role in decision making but that this significance is dependent on the legal rule that applies to the use of such materials. It indicates that the hierarchy debate is relevant and has potentially important implications on the results of the interpretive process. However, it is important to stress that the influence of preparatory work is not expected to be as strong in actual cases. In many cases, the difference between the preparatory work and the text is less clear than in the scenarios we used in this study (although two of them are based on actual cases). Moreover, in actual cases, it is expected that more relevant information will be available for the interpreters, which will influence the interpretation in addition to the text and the preparatory work.

More generally, our study suggests that the law matters to international law experts. It suggests that the rules of interpretation are important. As mentioned, the literature has questioned the effectiveness of the VCLT rules in the interpretive process, whereas our results suggest that the rules can actually affect treaty interpretation. In the current literature, insights taken from external theories on the legalistic model of decision making seem, in many cases, to become the starting point of the discussion. Our study emphasizes that for international law experts even secondary decision-making rules can play a significant role in their decision making. At least in some areas, experts can still resist external information and follow the legal norms as they understand them.[73]

In addition, our study highlights the importance of legal expertise: the unique role of the legal expert in interpreting the law. The results of our study are relevant

[70] See *supra* note 5.

[71] See Dotan, 'In Defence of Expansive Interpretation in the European Court of Human Rights', 3 *Cambridge Journal of International and Comparative Law* (2014) 508, at 516; Pauwelyn and Elsig, *supra* note 7, at 466, 468–469; Cohen and Kremnitzer, 'Judicial Activism: A Multidimensional Model', 18 *Canadian Journal of Law and Jurisprudence* (2005) 333, at 341.

[72] See Tobin, 'Seeking to Persuade: A Constructive Approach to Human Rights Treaty Interpretation', 23 *Harvard Human Rights Journal* (2010) 201, at 202.

[73] In a recent study, Spamann and Klohn suggested that the law might be less influential in judicial decision making than it is thought to be. See Spamann and Klohn, 'Justice Is Less Blind, and Less Legalistic, Than We Thought: Evidence from an Experiment with Real Judges' 45 *Journal of Legal Studies* (2016), 255. In contrast to our study, they examined the influence of a weak, non-binding, precedent on the result of an international criminal case and recognize that their findings do not necessarily apply to instances of strong precedents or clear law (at 273).

to the discussion of the importance of expertise in the behavioural research literature. As mentioned, using law students as proxies for judges is an accepted practice in behavioural research.[74] Studies have shown that experts are affected by biases and heuristics in many cases. Specifically, in the case of the effect of exposure to inadmissible evidence, research has shown that judges are often unable to ignore inadmissible evidence, just like lay persons.[75] Nonetheless, research on judges has found some instances where the judges were able to ignore inadmissible information.[76] For example, the prevalence of probable cause cases might explain the ability of judges to ignore information since they are able to use prior knowledge of similar cases.[77]

The prior familiarity explanation does not apply to our study since the scenarios we used do not necessarily represent common cases for all international law experts. Indeed, only a small minority of the IHRL experts in our study had prior familiarity with the *Litwa* case. Therefore, we offer a different possible explanation of the results. Legal interpretation is a unique legal expertise. Schauer has suggested that judges might act differently when it comes to tasks that are unique to their expertise. He specifically refers, *inter alia*, to legal interpretation in contrast to the task of fact-finding.[78] This observation is relevant not only to judges but also to any legal experts for whom legal interpretation is an integral part of their occupation. The results of the present study suggest that even law students in their advanced years might not be good proxies for legal experts when it comes to tasks that involve the latter's unique expertise, such as treaty interpretation.[79]

B *Limitations*

Our study is the first one to research the ability to ignore information in the context of the rules of treaty interpretation. In contrast to the many studies on inadmissible evidence, the present study is a first step in behavioural research on treaty interpretation. Further research is needed to strengthen and expand our findings. Second, most of the experts in this study had academic or practical expertise that did not include a judicial or semi-judicial role, while only a small percentage of the participants had served as judges or arbitrators. Nonetheless, the task of treaty interpretation is not unique to judges but is shared by all actors in the international law community. The lack of an institutional interpretive hierarchy in international law decreases the importance of courts, compared to their role in domestic law. Therefore, the participants in the experts study did not necessarily serve as proxies for judges but, rather, as part of the population for whom the hierarchy debate is

[74] See note 55 above.

[75] Wistrich *et al.*, *supra* note 16.

[76] *Ibid.*; Rachlinsli *et al.*, *supra* note 18.

[77] *Ibid.*, at 97.

[78] Schauer, *supra* note 20; compare Teichman and Zamir, *supra* note 12, at 692–693.

[79] These results are in line with the Kahan *et al.* study on motivated reasoning and professional judgment, which demonstrate that judges and legal experts, in contrast to law students, were found to be resistant to politically motivated reasoning when they performed tasks that were unique to their expertise. Kahan *et al.*, *supra* note 20, at 411–413.

relevant.[80] Moreover, our sample shares some of the main demographic variables of international judges: they are mostly European men with either an academic or civil service career and with substantial experience in international law. These qualifications seem to make the participants reasonable proxies for international judges.[81]

Third, in many cases, the decisions of international tribunals are made by a panel of judges. Empirical research has demonstrated that judicial decision making may be different in a panel setting than in individual judges' courtrooms.[82] In addition, group decision making has been found to moderate the influence of inadmissible evidence.[83] In our study, the participants made their decisions individually. In light of the results of the experts study, the effect of a panel decision would not be expected to be significant since experts seem to be resistant to the effect of exposure. However, it is possible to argue that group deliberation potentially influences experts' positions on the hierarchy debate or experts' decisions on the ambiguities of a specific case, similar to the way that the ideology of some judicial panel members has been found to influence the decisions of other panel members.[84] Future research should address these issues.

Fourth, as mentioned, the experiments on the students and the experts were conducted in different settings and different languages. A combined experiment on students and experts could strengthen our results on the qualitative difference between the two groups in their ability to ignore information. Finally, the study is an experimental study and does not involve real-world decision making. It is possible to suggest that when facing real consequences, interpreters tend to take preparatory work into account more than in cases where there is no actual cost for their decision. External validity is indeed a valid concern that is inherent in experimental studies.[85] It is important to note that this concern is mostly relevant to those who explain the difficulty of ignoring inadmissible evidence as motivationally based.[86]

5 Conclusion

Our study has focused on the possible effect of exposure to preparatory work of a treaty under the VCLT rules on treaty interpretation. Our results demonstrate that in

[80] See *supra* note 54.

[81] See D. Terris, C.P.R. Romano and L. Swigart, *The International Judge: An Introduction to the Men and Women Who Decide the World Cases* (2007), at 17–21. The main difference is the lack of domestic judges in our sample, which constitutes in the book's data around a third of the professional background of international judges.

[82] See, e.g., Teichman and Zamir, *supra* note 12, at 688–690; Cox and Miles, 'Judging the Voting Rights Act', 108 *Columbia Law Review* (*CLR*) (2008) 1; Choi and Gulati, 'Trading Votes for Reasoning: Covering in Judicial Opinions', 81 *South California Law Review* (2008) 735.

[83] London and Nunez, 'The Effect of Jury Deliberations on Jurors' Propensity to Disregard Inadmissible Evidence', 85 *Journal of Applied Psychology* (2000) 932.

[84] See Cox and Miles, *supra* note 90.

[85] See, e.g., Chilton and Tingley, 'Why the Study of International Law Needs Experiments', 52 *CJTL* (2013) 173, at 231.

[86] Wistrich *et al.*, e.g., suggest that policy preferences have a limited role in the difficulty to ignore information. See Wistrich *et al.*, *supra* note 16, at 1323.

contrast to law students, international law experts are able to resist the effect of exposure to preparatory work when it is prohibited to use such material. The article stresses the importance of the VCLT rules and the role of the legal expert in the interpretive process. It also provides a window on the positions prevalent among international law experts on the hierarchy of the VCLT rules. Moreover, it shows the uniqueness of the IHRL community among the different interpretive communities in international law.

The study may have implications well beyond the realm of treaty interpretation. Indeed, the VCLT rules are quite unique, and many jurisdictions do not establish statutory rules of interpretation, let alone rules that mandate a strict hierarchy between different methods of interpretation. A notable example is the lack of federal rules of statutory interpretation in the USA.[87] Nonetheless, in some countries, similar rules of either statutory interpretation or contract interpretation have been adopted. Thus, for example, the Australian Acts Interpretation Act of 1901 uses wording that is almost identical to Article 32 when it addresses the circumstances that allow the use of extrinsic materials in the interpretation of an act.[88] In addition, even in systems that lack clear rules of interpretation, legal interpretation attracts much attention, and the potential establishment of rules of interpretation is an important part of the discussion.[89] Legislative history and circumstances that are external to the text play an important role in the discussions on statutory and contract interpretation.[90] Our study sheds light on the potential effect of exposure to such materials, which should be taken into account in the theoretical discussion of the issue.

It may be argued that the ability to generalize our conclusions in the realm of domestic law is limited as a result of the relevant differences between international and domestic law experts rather than their general expertise in legal interpretation. Thus, for example, it is possible to suggest that international law experts will follow the rules of interpretation more strictly than domestic law experts, since the long-standing questioning of the legitimacy of international law as 'real law' results in a tendency of international law experts to be more formalistic in their approach to legal norms than their domestic counterparts.[91] Further research should be conducted on domestic law experts in order to examine whether the tendency to resist the influence of the exposure to information applies to them as well.

[87] See, e.g., Staszewski, 'The Dumbing Down of Statutory Interpretation', 95 *Boston University Law Review* (2015) 209, at 211–212.

[88] *Acts Interpretation Act* 1901 (Cth), s. 15AB(1) (Australia); Restatement (Second) of Contracts §§ 200–223 (1981); Zamir, 'The Inverted Hierarchy of Contract Interpretation and Supplementation', 97 *CLR* (1997) 1710, at 1716, n. 2 (presenting hierarchies in contract interpretation in several legal systems).

[89] See, e.g., O'Connor, 'Restatement (First) of Statutory Interpretation', 7 *New York University Journal of Legislation and Public Policy* (2004) 333; Rosenkranz, 'Federal Rules of Statutory Interpretation', 115 *Harvard Law Review* (2002) 2085.

[90] See, e.g., Fleischer, 'Comparative Approaches to the Use of Legislative History in Statutory Interpretation', 60 *American Journal of Comparative Law* (2012) 401; Law and Zaring, 'Law versus Ideology: The Supreme Court and the Use of Legislative History', 51 *William and Mary Law Review* (2010) 1653.

[91] For the importance of the rules of interpretation for the international lawyer in this regard, see Klabbers, *supra* note 2, at 276–277. For another account of the potential influence of the need to defend international law on international lawyers, see Koskenniemi and Leino, 'Fragmentation of International Law? Postmodern Anxieties', 15 *Leiden Journal of International Law* (2002), 553, at 558.

Appendix 1: Scenarios[92]

Detention of a Drunk Man

Imagine that you are a judge at the European Court of Human Rights (ECtHR). The case concerns a person named Witold Litwa. On May 2014, Litwa drank a couple of glasses of an alcoholic beverage for the first time in his life. He then appeared drunk at a post office in Krakow and insulted the postal service workers. Police officers who arrived at the scene believed Litwa that this was the first time he had drunk alcohol, but nevertheless took him and held him for 12 hours in detention. He was then released.

Article 5(1)(e) of the European Convention on Human Rights (ECHR) allows detention in the following circumstances: (e) the lawful detention of persons for the prevention of the spreading of infectious diseases, alcoholics or drug addicts or vagrants.

(*Treatment*: The following is part of preparatory work of the ECHR, which appears in the statement of reasons, incorporated in the *Report of the Committee of Experts*, which describes the purpose of Article 5(1)(e): 'The article's purpose is to protect public order. Its text covers, specifically, the right of signatory states to take necessary measures for combating disturbances to public order as a result of public drunkenness.')

Poland argues that his arrest was legal under Article 5(1)(e) of the ECHR, which includes, in their view, arrest in cases of public drunkenness. Litwa argues that his arrest was unlawful since he was drunk for the first time and not an alcoholic, as the article requires.

Was it permissible to arrest Litwa under Article 5(1)(e)? Yes / No

Deportation (translated from Hebrew)

Imagine that you are adjudicating a case in which Israel wishes to deport a senior Hamas member, who was involved in the planning of terror attacks against civilians, outside the occupied territories. His lawyers present Article 49(1) of Geneva Convention IV, which reads as follows:

> Individual or mass forcible transfers, as well as deportations of protected persons from occupied territory to the territory of the Occupying Power or to that of any other country, occupied or not, are prohibited, regardless of their motive.

(*Treatment*: Below is a description of the background of the creation of Article 49(1) of Geneva Convention IV as it is described in the preparatory work of the treaty:

> The article is the result of the Second World War. Article 49(1) is the most important subparagraph of article 49, in that it prohibits the forcible transfer or deportation of persons from occupied territories. There is doubtless no need to give an account here of the painful recollections called forth by the 'deportations' of the Second World War, for they are still present in everyone's memory. It will suffice to mention that millions of human beings were torn from their homes, separated from their families and deported from their country, usually under

[92] We present the second experts' version of the detention and bilateral investment treaty scenarios and a translated version of the deportation scenario.

> inhumane conditions. These mass transfers took place for the greatest possible variety of reasons, mainly as a consequence of the formation of a forced labor service. The thought of the physical and mental suffering endured by these 'displaced [p. 279] persons', among whom there were a great many women, children, old people and sick, can only lead to thankfulness for the prohibition embodied in this paragraph, which is intended to forbid such hateful practices for all time.)

Does the deportation of the Hamas member outside the occupied territories violate Article 49(1) of Geneva Convention IV? Yes / No

Bilateral Investment Treaty Dispute

In early 2010, GGL, a Slovenian company, entered into a contract with the government of Sweden 'for the acquisition, management, operation and disposition of the Kallak Iron Mine'. According to the contract, the Swedish government was obligated to transfer all the shares of the state-owned company of the mine immediately after the transfer of the agreed payment. However, the government transferred the shares a full year and a half after the payment was made. On June 2004, the Agreement between the Government of the Republic of Slovenia and the Government of the Kingdom of Sweden on the Promotion and Mutual Protection of Investment entered into force.

The bilateral investment treaty (BIT) includes, *inter alia*, the following articles:

> Article 8(3) – 'Each Contracting Party shall observe any other obligation it has assumed with regard to investments in its territory by nationals or companies of the other Contracting Party.'
> Article 20 – 'The investor shall be entitled to submit any dispute premised on this convention to the International Centre for the Settlement of Investment Disputes (ICSID).'

GGL submitted a request for arbitration to ICSID, arguing that the delay in the transfer of shares constituted a breach of the contract and that such a contract breach also violated Article 8(3) of the BIT. The Swedish government agreed that the contract was violated but argued that the BIT was not intended to address mere commercial contractual disputes, and, therefore, Article 8(3) did not cover such contract breaches. Both sides agreed that no other article of the BIT was violated.

(*Treatment*: In the draft of the BIT, before the adoption of the final version, Article 8(3) appeared: 'Each Contracting Party shall observe any other obligation it has assumed with regard to investments in its territory by nationals or companies of the other Contracting Party, including any contractual obligations between the contracting party and nationals or companies of the other contracting party.' The protocol of the negotiations reveals that the head of the Slovenian delegation stated that Slovenia 'objects to this wording since regular contractual obligations, in contrast to obligations which reflect the role of the state as a sovereign, should not be part of the convention protection.' Following this statement, and a short discussion between the representatives of the two countries, Sweden agreed to change the text of Article 8(3) to its current form, so that it would be in line with the Slovenian position.)

Does the alleged contractual breach violate article 8(3) of the BIT? Yes / No

Appendix 2: The VCLT Rules of Treaty Interpretation in the Second Students Experiment (translated)

'Rules allow' version of the VCLT rules:

31. General Rule of Interpretation

(1) A treaty shall be interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty in their context and in the light of its object and purpose.
(2) The context for the purpose of the interpretation of a treaty shall comprise the text, including its preamble and annexes, the preparatory work of the treaty and the circumstances of its conclusion.

'Rules Prohibit' version of the VCLT rules:

31. General Rule of Interpretation

(1) A treaty shall be interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty in their context and in the light of its object and purpose.
(2) The context for the purpose of the interpretation of a treaty shall comprise the text, including its preamble and annexes.

32. Supplementary Means of Interpretation
It is not allowed to use supplementary means of interpretation including the preparatory work of the treaty and the circumstances of its conclusion.