Canadian Journal of Law & Jurisprudence XXXIII No.2 August 2020, 435-460
© The Author(s) 2020. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.
doi: 10.1017/cjlj.2020.5

Categories, Balancing, and Fake News: The Jurisprudence of the European Court of Human Rights

Alessio Sardo

1. Introduction

Freedom of expression has been often described as a necessary precondition for democracy and for the implementation of an effective system of human rights. A deliberative democracy cannot function if citizens are not granted the fundamental right to express their views and to criticize the government without being censored.[1] The rule of law becomes an empty notion if legal orders do not protect the impartial, autonomous judgments of the judiciary.[2]

Akhil R. Amar argued that free speech—the paradigmatic case of freedom of expression—existed even before the First Amendment.[3] The very act of enacting the U.S. Constitution presupposed, so the argument goes, that the constituent power was already granted with extensive freedom of speech during the process of adoption of the Federal Constitution. This freedom was rooted in the normative practices of "We, the People." In Europe, Robert Alexy pointed out that the ultimate justification for the existence of human rights beyond positive law stems from the very possibility of a universal and institutionalized discourse where

---

This essay is a thoroughly revised and expanded version of the paper delivered at the *Humboldt Kolleg* "Freedom of Expression and the Right of Privacy in the Jurisprudence of the ECHR" (Strasbourg, 26-27 April 2019), in the session for young researchers. The Author wishes to extend his gratitude to Mauro Barberis, Martin Borowski, and Peggy Ducoulombier for rewarding discussion of the issues. A special thanks goes to Alex Alferiev of CJLJ, too, for her fine editorial work and suggestions. The Author remains grateful to the Alexander von Humboldt Foundation for sponsoring his research.

Alessio Sardo, Alexander von Humboldt Fellow, University of Heidelberg, Department of Public Law, Constitutional Theory and Philosophy of Law. alessiosardo@ymail.com

1. The ECHR has explicitly endorsed this claim. See *Lingens v Austria*, No 9815/82 (1986) 8 EHRR 407 at paras 42, 44, 45 [*Lingens*]; *Şener v Turkey* (2000), Eur Ct HR 377 at para 40 [*Şener*]; *Thoma v Luxembourg*, No 38432/97, [2001] III ECHR 41; *Marônek v Slovakia*, No 32686/9619, [2001] III ECHR 337 at para 52; *Dichand and Others v Austria*, No 29271/9526, (February 2002) at para 37. See also *Handyside v the United Kingdom*, No 5493/72, (1976) 1 EHRR 737 at para 49 [*Handyside*].
2. Jochen A Frowein, "Freedom of Expression Under the European Convention on Human Rights" (2004) 97 (3) Monitor/Inf 1.
3. Akhil R Amar, "How America's Constitution Affirmed Freedom of Speech Even Before the First Amendment" (2010) 38 Cap U L Rev 502 at 513. On the role that freedom of expression plays in spreading a liberal form of constitutionalism see Michael Ronsenfeld & András Sajó, "Spreading Liberal Constitutionalism: An Inquiry into the Fate of Free Speech in New Democracies" in Sujit Choudhry, ed, *The Migration of Constitutional Ideas* (Cambridge University Press, 2006) at 146 and Lee C Bollinger, *The Tolerant Society: Freedom of Speech and Extremist Speech in America* (Oxford University Press, 1986).

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

every participant can ask for and offer reasons.[4] While Amar relies on a structural argument with a 'historical twist', Alexy makes a more philosophical point. Notwithstanding this difference, the central claim raised by the two leading constitutional scholars is the same: the institutional act of enacting a constitution *presupposes* freedom of expression.

It would be tempting to conclude that freedom of expression is the most fundamental right, not from an axiological perspective (namely, the moral ranking of human and fundamental rights), but from a purely justificatory and conceptual standpoint.[5] Is it not the case that the European Court of Human Rights (hereinafter, ECtHR) emphasized that freedom of expression constitutes "one of the essential foundations" for the existence and implementation of any other right of the Convention and "the development of every man"?[6]

The main goal of this essay is to explore the methods used by the ECtHR for balancing freedom of expression against competing rights, with an eye to the current debate on fake news. Section 2 offers a survey of the normative framework of Article 10 of the European Convention of Human Rights (ECHR). Section 3 explains how the Court's method for weighing freedom of expression against competing rights is much more structured than an *ad hoc* balancing because it is based on *categorization* and *sliding-scales*. Section 4 deals with the challenge of fake news. Taking cues from the case law and borrowing from recent developments in philosophy, a *minimal definition* of fake news will be proposed as a safeguard against censorship in disguise, "collateral" over-censorship, and judicial arbitrariness. Definitions that are too broad, vague, and ambiguous are open to abuse, and this could jeopardize freedom of expression. The present inquiry will also disentangle the relevant factors for balancing web-expression and isolate

---

4. See Robert Alexy, "Law, Morality, and the Existence of Human Rights" (2012) 25(1) Ratio Juris 2. A similar argument was proposed by Jürgen Habermas, *Communication and the Evolution of Society* (Beacon Press, 1979) at 2-4, 32.

5. For an argument on why expression is fundamental for our self-realization see Jonathan Gilmore, "Expression as Realization: Speaker's Interest in Freedom of Speech" (2011) 30(5) Law & Phil 517. For a sociological inquiry on this issue see James W Nickel, "Freedom of Expression in a Pluralistic Society" (1988/89) 7(3) Law & Phil 281. Amartya Sen claims that there has never existed a famine system with free press and democratic elections. See Amartya Sen, *Poverty and Famines* (Oxford University Press, 1981). See also Amartya Sen, *Development as Freedom* (Knopf, 1999).

6. *Handyside*, *supra* note 1 at para 49; *Vereinigung Bildender Künstler v Austria*, No 68354/01 [2007] ECHR 79. For a general analysis of the European Convention of Human Rights (ECHR), which was signed in Rome on 4 November 1950, see Martyn Bond, *An Introduction to the European Convention on Human Rights* (Council of Europe Publishing, 2018); ECHR, *50 Years of Activity. The European Court of Human Rights. Some Facts and Figures* (Council of Europe Publishing, 2010); David J Harris et al, *Law of the European Convention on Human Rights* (Oxford University Press, 2009). It is noted that the precedents of the European Court of Human Rights (ECHR) have a binding force on Member States. See Anja Sieber-Fohr & Mark E Villiger, eds, *Judgments of the European Court of Human Rights—Effects and Implementation* (Nomos, 2014). The ECHR has the final word on the interpretation of the text of the Convention; however, the Court can be accessed only as a last resort, when all internal remedies have run out. The so-called doctrine of the *domestic margin of appreciation* plays a fundamental role for the vertical separation of powers between ECHR and Member States by establishing the degree of deference that might be granted to national governments.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

several sub-categories of fake news. The final section contains a summary of the arguments and a denouement.

2. Freedom of Expression in the Balance

Defining the contours of freedom of expression is not an easy task. A first major issue is the risk of conflating freedom of expression with a general right to liberty or autonomy;[7] a second puzzle concerns judicial balancing and the idea that freedom of expression is not an absolute right.[8]

The ECtHR, on the one hand, endorses a broad reading of the notion of expression, and, on the other, relies on balancing *qua* its methodology for introducing some exceptions outside the core of this human right. The Court's approach finds textual support in Art 10 of the ECHR, which reads as follows:

1. Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This Article shall not prevent States from requiring the licensing of broadcasting, television or cinema enterprises.
2. The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary.

However, the Court's interpretation of Art 10 goes beyond its plain meaning: freedom of expression is not limited to the liberty of holding, sharing, receiving information and ideas *through speech acts*.[9] It also comprises (i) the freedom to produce, exhibit, and distribute artworks; ii) the liberty of realizing a public artistic performance;[10] iii) the access to relevant public information;[11] iv) the liberty of displaying pictures, icons, images,[12] and religious or political symbols;[13] v) preserving cultural heritage,[14] broadcasting,[15] and satirical

7. Frederick Schauer, *Free Speech: A Philosophical Enquiry* (Cambridge University Press, 1982) at 5ff.
8. Jean-François Flauss, "The European Court of Human Rights and the Freedom of Expression" (2009) 84(3) Ind LJ 809 at 813.
9. John L Austin, *How to Do Things with Words* (Clarendon Press, 1962); John Searle, *Speech Acts* (Cambridge University Press, 1969).
10. See *Müller and Others v Switzerland* (1988), 13 EHRR 212.
11. See *Leander v Sweden* (1987), 9 EHRR 433.
12. See *Chorherr v Austria* (1994), 17 EHRR 358.
13. See *Vajnai v Hungary*, No 33629/06, [2008] IV ECHR 173.
14. See *Stevens v the United Kingdom*, No 11674/85 (3 March 1986).
15. See *Groppera Radio AG and Others v Switzerland* (1990), 12 EHRR 321.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

expression;[16] vi) the protection against self-incrimination.[17] The sphere of protection granted to the freedom of the press and political speech is even broader: it includes information and ideas that "offend, shock or disturb."[18] Political speech, too, is highly protected: for instance, obscene protests involving the sculpture of a large penis, held in front of the Protector's General Office,[19] or extremist political discourse are permitted under the ECHR legal regime.[20] Furthermore, if applied to the press, the restrictions established by the Second Paragraph of Art 10 must be "construed narrowly."[21] A peculiarity of the ECHR system is the protection of the form in which ideas and information are conveyed, also when the tone is polemical and even aggressive.[22]

The Court allows the possibility of restricting expression by granting priority to a Contracting State's interests (national security, public safety, and protection of health and morals) as well as to other, conflicting, human rights such as the right to privacy (Art 8 ECHR).[23] In case of conflicts between rights (or between a right and a State's interest) the Court *strikes a balance*, deciding which right or interest takes priority over the other, and specifying under which conditions the relation of precedence holds.

In shaping the relations of conditional priority, the Court has always relied on an evolutionary approach, based on a form of "living constitutionalism."[24] The ECtHR does not interpret Art 10 according to its original meaning because the construction Art 10 is always the result of a "dynamic update" which takes into account societal values and beliefs, both from the internal perspective (the position of the participant) and from the external perspective (looking at national systems from a broader, European, standpoint).[25]

3. More than a Simple *Ad Hoc* Balancing

The Court's method for weighing principles is much more sophisticated, structured, and responsive than a sheer *ad hoc*, freewheeling, balancing of competing

16. See *Eon v France*, No 26118/10, [2013] ECHR 216.
17. See *K v Austria* (1992), 62 Eur Comm'n HR DR 216.
18. See *Axel Springer v Germany*, No 39954/08, [2012] ECHR 227.
19. See *M t saru v the Republic of Moldova*, Nos 69714/16 and 71685/16 (15 January 2019).
20. See *Stomakhin v Russia*, No 52273/07, [2018] ECHR 390.
21. See *Cox v Turkey*, No 2933/03 at para 31 (20 May 2010).
22. See, e.g., *De Haes and Gijsels v Belgium* (1997), 25 EHRR 1 at para 48; *News Verlags GmbH & Co KG v Austria*, No 31457/96, [2002] ECHR 96 at para 39; *Campos Dâmaso v Portugal*, No 17107/05 (24 April 2008) at para 31.
23. The most common forms of restriction and interference with freedom of expression include criminal convictions, censorship, orders to pay damages, the prohibition to exercise the profession of journalism, the refusal to grant a broadcasting license, and denying the disclosure of information —however, this is an open list. See, e.g., *The Sunday Times v the United Kingdom* (1979), 2 EHRR 245 [*Sunday Times*]; *Observer and Guardian v the United Kingdom* (1991), 14 EHRR 153 [*Observer and Guardian*]; *Frankowicz v Poland*, No 53025/99, (16 December 2008).
24. David A Strauss, *The Living Constitution* (Oxford University Press, 2010).
25. On the distinction between the internal and external point of view see HLA Hart, *The Concept of Law*, 2d ed (Clarendon Press, 1994) at 91ff and Robert Alexy, *The Argument from Injustice* (Clarendon Press, 2000) at 27-68.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

interests.[26] In fact, the jurisprudence of the ECtHR combines two different approaches: the *categorical approach*, which establishes what counts as freedom of expression through stipulations, and the *balancing approach*, which proceeds through a case-sensitive assessment of the factual elements within previously isolated categories. This combination can be understood as an attempt in rational systematization: the weighing of competing rights and interests is performed *through* categories,[27] and the Court gradually develops clusters of rules establishing different *levels of scrutiny* and *burdens of justification* that the arguments in favor of the challenged law or measure must satisfy. These clusters of rules are "activated" by the type of expression at hand. For instance, only a light burden of justification is usually required for measures that interfere with commercial expression,[28] whereas a higher threshold must be met for justifying a gag-order imposed upon the press.[29] In this respect, the ECtHR follows a jurisprudential pattern which is quite close to the one endorsed by the U.S. Supreme Court. The ECtHR, too, has gradually developed a full array of categories of expression and balancing tests, which operate in virtually every corner of the ECHR system.[30]

On many occasions, the ECtHR has relied on *proportionality* as a method for resolving conflicts between freedom of expression and competing rights. A disclaimer is in order here: the term 'proportionality' has two different meanings, *proportionality as a balancing test* and *proportionality as a (meta-)principle that underpins every operation of balancing*. In the first sense, proportionality denotes a three-pronged method for resolving a conflict between rights: the first phase checks the suitability of the measure; the second phase ascertains the necessity of the measure; the final step settles the relation of conditional precedence

26. Laurent B Frantz, "Is the First Amendment Law: A Reply to Professor Mendelson" (1963) 51(4) Cal L Rev 729 at 732ff. Frantz asks whether a case-by-case approach to balancing is vital for freedom of expression, or whether it is possible to establish rules for whole classes of cases.

27. On categories and balancing see Joseph Blocher, "Categoricalism and Balancing in First and Second Amendment Analysis" (2009) 84(2) NYUL Rev 375; Daniel A Farber, "The Categorical Approach to Protecting Speech in American Constitutional Law" (2009) 84(3) Ind LJ 917; Elizabeth Zoller, "The United States Supreme Court and the Freedom of Expression" (2009) 84(3) Ind LJ at 885. On the combination of the categorial approach and balancing in the ECHR jurisprudence, see Peggy Ducoulombier, "Conflicts between Fundamental Rights and the European Court of Human Rights" in Eva Brems, ed, *Conflicts between Fundamental Rights* (Antwerp: Intersentia, 2008) at 217. See also Samantha Besson, "Human Rights in Relation" in Stijn Smeth & Eva Brems, eds, *When Human Rights Clash at the European Court of Human Rights: Conflict or Harmony?* (Oxford University Press, 2017) at 23. Besson uses the term "qualitative balancing" for a third way between *ad hoc* balancing and categorization.

28. The Court has carved the notion of 'commercial expression' in a series of precedents. See, e.g., *Casado Coca v Spain* (1994), 18 EHRR 1; *Hertel v Switzerland* (1998), 28 EHRR 534 [*Hertel*]; *Markt Intern v Germany* (1989), 12 EHRR 161.

29. *Girleanu v Romania*, No 50376/09 (26 September 2018) at paras 68-98.

30. See Harry Kalven Jr, *A Worthy Tradition. Freedom of Speech in America* (Harper & Row, 1988); Frederick Schauer, *The Law of Obscenity* (The Bureau of National Affairs Inc, 1976) at 116-36. Alexander T Aleinikoff, "Constitutional Law in the Age of Balancing" (1987) 96 Yale LJ 943 at 966-72; Kate Klonick, "The New Governors: The People, Rules, and Processes Governing Online Speech" (2018) 131 Harv L Rev 1598 at 1658-62.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

between the colliding rights or principles.[31] In the second sense, proportionality is understood as a broad principle which prescribes that any interference with a human right must be commensurate for the legitimate objective pursued by the legislator. While the proportionality test is used only for specific categories of speech, proportionality *qua* meta-principle orientates every single balancing operation of the ECtHR.

The Court has relied on the three-pronged proportionality test for resolving conflicts between freedom of expression and national security,[32] or between freedom of expression and liberty rights.[33] The proportionality test is frequently employed also for conflicts between Art 10 and Art 8: it is coupled with strict scrutiny for those classes of cases requiring the highest protection. For instance, in *Observer and Guardian v. United Kingdom*, the Court held that "the dangers inherent in prior restraints are such that they call for the most careful scrutiny on the part of the Court. This is especially so as far as the press is concerned, for news is a perishable commodity and to delay its publication, even for a short period, may well deprive it of all its value and interest."[34]

The proportionality test is ubiquitous,[35] but the most common balancing method is an altogether different kind of three-pronged test which we might label "the necessity test."[36] The three-pronged necessity test finds textual support in Art 10(2), and requires the fulfillment of three conditions for a justified interference with the freedom of expression.[37] *First*, the interference must be prescribed

---

31. Robert Alexy, "Constitutional Rights and Proportionality" (2014) 22 Revus 51; Ernst Wolfgang Böckenförde, *Constitutional and Political Theory: Selected Writings* (Oxford University Press, 2017) at 254; Martin Borowski, *Grundrechte als Prinzipien* (Nomos, 2018) at 254-72. There is a lively, on-going debate over the impact of proportionality on the application of human rights. See, e.g., Stavros Tsakyrakis, "Proportionality: An Assault on Human Rights?" (2009) 7 Intl J Const L 468; Matthias Klatt & Moritz Meister, "Proportionality: A Benefit to Human Rights?" (2012) 10 Intl J Const L 687.

32. *Hadjianastassiou v Greece* (1992), 16 EHRR 219 at paras 45-47; *Stoll v Switzerland* (2007), 44 EHRR 53 at para 151.

33. *Open Door and Dublin Well Woman v Ireland* (1992), 15 EHRR 244. The Court held that a total ban on the distribution of brochures informing pregnant women about the possibilities of abortion abroad was *disproportionate* despite the strong commitment to Catholic values of the Irish Constitution.

34. *Observer and Guardian*, *supra* note 23 at para 60.

35. *Gawęda v Poland*, No 26229/95, [2002] ECHR 95; *Sunday Times*, *supra* note 23; *Observer and Guardian*, *supra* note 23.

36. Flauss, *supra* note 8 at 813: "Serving as standards, or as a form of control of the 'necessity' of restrictive measures introduced or tolerated by states relative to the freedom of expression, 'discretionary powers' are by definition a means of favoring—in casu (in case of extreme necessity)—the Strasbourg Court's freedom of assessment and, consequently, authorizing its possible manipulation. [. . .] The multiplicity of variables analyzed in order to balance restrictions placed on the freedom of expression effectively offers the European judge wide powers of flexibility in terms of the interests at issue, especially within the framework of a jurisprudential and jurisdictional policy based on the 'special case'."

37. See Dominika Bychawska-Siniarska, *Protecting the Right to Freedom of Expression Under the European Convention of Human Rights: A Handbook for Legal Practitioners* (Council of Europe, 2017) at 32ff; Dirk Voorhoof, "The European Convention on Human Rights: The Right to Freedom of Expression and Information restricted by Duties and Responsibilities in a Democratic Society" (2015) 7(2) Human Rights 1 at 6ff.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

by foreseeable, precise, and publicly accessible law;[38] *second*, the interference shall aim to protect significant interests (territorial integrity, public safety, prevention of disorder and crime, protection of health and reputation, preventing the disclosure of information received in confidence, and maintaining the authority or impartiality of the judiciary); *third*, the interference must be *necessary* for preserving the democratic values. Determining what counts as 'necessary'—which is crucial for both the proportionality test and the three-pronged necessity test—is one of the most difficult tasks that the ECtHR has to fulfill. As explained in *Observer and The Guardian v. the United Kingdom*, the Court considers that a measure is necessary only if a "pressing social need" exists.[39]

A third test, a form of rational basis "with bite", is confined to cases involving commercial expression.[40] It is the least protective test among the ones used for Art 10. The rational basis with bite considers the appropriate measures that suitably pursue the goal(s) of the legislature, according to a means/ends analysis. The test is yoked to a high standard of deference towards the Member States, which usually grants them a wide margin of appreciation. Under extenuating conditions, a national government may even impede the publication of the exact description of a publicly relevant fact.[41] However, contracting states are not entitled to act arbitrarily.[42] They are required to pursue one of the legitimate ends mentioned in the second paragraph of Art 10, which, in turn, the Court understands as indefeasible (i.e. no implicit exception is allowed). The metaphor of the "bite" is here used precisely for grasping the resistance to implicit exceptions, which strengthens the means/ends rationality requirement. The rational basis test, too, is coupled with the duty to *rule through laws* meaning that every interference with freedom of expression must be grounded in legal provisions.[43]

The ECtHR has also designed more detailed tests for balancing freedom of expression in three recent landmark decisions: *Axel Springer v. Germany*,[44] *von Hannover v. Germany*,[45] and *Perinçek v. Switzerland*.[46] These are all examples of *definitional balancing*: the Honorable Judges have relied on categories of expressions, a set of procedural rules orienting the deliberative process, and specific burdens of justification designed for a specific class of cases.

The judgment of *Axel Springer* concerns the publication of two articles revealing that Mr. X, a well-known actor who played the role of Police Superintendent Y in a television series, had been convicted of unlawful possession of drugs. The

---

38. See, e.g., *Editorial Board of Pravoye Delo and Shtekel v Ukraine*, No 33014/05, [2011] ECHR 5 [*Shtekel*]; *Youth Initiative for Human Rights v Serbia*, No 48135/06 (25 June 2016).

39. *Observer and Guardian*, *supra* note 23 at para 59; *Hertel*, *supra* note 28 at para 46.

40. Gayle L Pettinga, "Rational Basis With Bite: Intermediate Scrutiny by Any Other Name" (1987) 62(3) Ind LJ 779 at 780ff.

41. *Markt intern Verlag GmbH and Klaus Beerman v Germany* (1989), 12 EHRR 161 at para 35.

42. Frederick Schauer, "Giving Reasons" (1995) 47(4) Stan L Rev 633.

43. The Court has introduced a notable exception for the British common law system. See *Sunday Times*, *supra* note 23.

44. *Axel Springer*, *supra* note 18.

45. Nos 40660/08 and 60641/08, [2012] ECHR 835.

46. No 27510/08, [2015] VI ECHR 181.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

German newspaper *Bild* published these articles under satirical titles, accompanied by three pictures of the TV star. In order to reach a decision, the ECtHR relied on a multi-pronged test based on the following criteria: i) the general interest of the debate; ii) how well known the person involved is; iii) the subject of the report; iv) the prior standard of conduct of the person concerned; v) the method for obtaining the information and its veracity; and (vi) the consequences of publications. The case of *von Hannover* introduces analogous criteria for the publication of photographs: i) the degree of notoriety of the person affected; ii) the subject of the picture; iii) the prior conduct of the person involved; and (iv) the consequence caused by the publication of the photos. Finally, *Perinçek*—a well-known case concerning the denial of the Armenian Genocide—counts as an attempt in systematization too. The Court has analytically isolated eight relevant factors for evaluating cases involving hate speech: i) the nature of the statements; ii) the geographical, historical, and temporal context; iii) the extent to which the statement affects the competing rights; iv) the existence of a consensus; v) the possibility that the interference is prescribed by the international obligation assumed by the State; vi) the method employed by the State to justify the conviction; vii) the severity of the interference; and (viii) the final weighing of freedom of expression against the right to private life.

Coming full circle, a traditional form of *strict scrutiny* applies to the freedom of the press. As the Court emphasized in a leading case *Jersild v. Denmark*,[47] the "punishment of a journalist for assisting in the dissemination of statements made by another person in an interview would seriously hamper the contribution of the press to discussion of matters of public interest and should not be envisaged unless there are particularly strong reasons for doing so."[48] The press is not immune from liability, although this category "triggers" strict scrutiny, provided that the journalist is acting in good fate and her information is reliable. With an eye to potential conflicts with Art 8, the Court also clarifies that personal reputation prevails only when the attack to privacy is severe.[49]

With regard to context-sensitivity, the jurisprudence of the ECtHR follows *a sliding-scale approach*, similar to the one that was advocated by Justice Thurgood Marshall of the U.S. Supreme Court.[50] The term "sliding-scale" denotes a balancing approach that associates a *continuum* of levels of scrutiny with different categories, which can be adjusted to the peculiarities of the case at hand. A category, in turn, is a bundle of interrelated properties and principles that define, by denoting a class of objects, entities, and actions.[51] On a sliding scale, the operative degree of scrutiny depends on a mix of factors: the nature of the classification in question, the constitutional and societal importance of

47. (1994), 19 EHRR 1.

48. *Verlagsgruppe News GmbH v Austria*, Nos 76918/01 and 10520/02 (14 December 2006) at para 31; *Becker v Norway*, No 21272/12 (5 October 2017) at para 35.

49. *Axel Springer*, *supra* note 18 at para 83.

50. Glenn L Starks & Erik F Brooks, *Thurgood Marshall: A Biography* (Greenwood, 2012).

51. Frederick Schauer, "Categories and the First Amendment: A Play in Three Acts" (1981) 34 Vand L Rev 265 at 277ff.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

the individual interest adversely affected, and the strength of the state interest asserted in support of the law under consideration.[52]

In a nutshell, the ECtHR is not committed to *a priori* rankings between rights that consider freedom of expression as the "ace of trumps". The balancing tests endorsed by the Court, as well as the sliding-scale approach, are part and parcel of what we might call a "grand theory" for applying human rights.

As hinted above, the Court's approach involves a preliminary process of *categorization*.[53] The Court does not treat the ECHR as a system of absolute rights whose boundaries can be crafted with a precise, ostensive definition of which cases instantiate a particular right or an analytical definition of the semantic content of that particular right. Nevertheless, definitions map different categories of expression coupled with varying levels of protection and specific balancing tests.[54] The use of categories is not an attempt to purge balancing out of the ECHR system: definitions provide the scheme for allotting the appropriate burdens of justification and the appropriate tests. While judges and journalists are at the top of the scale, the categories of militaries and civil servants are placed right at the bottom.[55] Commercial expression, too, is with bite.

On many occasions, the Court has resorted on "content-based" categorization (namely, classifications based on the content). This type of categorization pursues two main goals: i) leaving certain forms of expression outside the protection of Art 10; ii) specifying sub-categories, in order to address them separately. The most striking example of the former strategy is Holocaust denial: a concept carefully situated outside the scope of Art 10.[56] An excellent example of the second strategy is offered by the four "maxi-categories" of speech: description of facts, value judgments, opinions, and insults. Value judgments, due to their intrinsic subjective nature, benefit extensive protection:[57] under the ECtHR's perspective, a value-judgment is nothing else than a positive or negative attitude towards a particular action, item, or state of affairs (i.e. being for/against). As for opinions and descriptions, the possibility of restrictive measures depends on the general

52. Jeffrey M Shaman, *Constitutional Interpretation. Illusion and Reality* (Greenwood, 2001) at 102.

53. On this specific point, the present inquiry disagrees with the analysis offered in Flauss, *supra* note 8 at 810. The opposition between the US tradition and the EU tradition is not as radical as it appears at first glance, and these similarities have a historical explanation. As Flauss himself notices: "The attachment to this ideological stance has grown stronger since the middle of the 1980s, notably with the presence inside the court of judges elected on account of "new European democracies." The judges, often having diplomas from American universities and sometimes even academic experience overseas, have promoted the legal precedents set by the United States Supreme Court relating to "freedom of expression" to a preferential source of inspiration" at 811.

54. Schauer, *Free Speech*, *supra* note 7 at 32f.

55. See *Castells v Spain* (1992), 14 EHRR 445; *Vereinigung demokratischer Soldaten Österreichs and Gubi v Austria* (1994), 20 EHRR 56 at para 38. See, generally, Dirk Voorhoof & Tarlach McGonagle, eds, *Freedom of Expression, the Media and Journalists: Case law of the European Court of Human Rights* (European Audiovisual Observatory, 2013).

56. *Garaudy v France*, No 65831/01, [2003] ECHR 1. For an insightful analysis of this strategy see Andras Sajó, ed, *Abuse: The Dark Side of Fundamental Rights* (Eleven International, 2006).

57. *Lepojic v Serbia*, No 13909/05 (6 November 2007).

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

interest for the public debate: "[t]he extent of the State's recognized power of restriction structurally dependents on the contribution the speech or message makes to a general interest or public debate. The existence of a general interest debate leads *ipso facto* to a strengthening of European control. The same goes for political discourse."[58]

Freedom of expression offers a reason for imposing waves of duties and responsibilities on those States that signed the ECHR.[59] However, categorization allows to isolate a "core" of the right, which includes the privilege to receive information and share ideas as well as the correlative duty of the State to facilitate the uptake of communication. The Member States shall not disrupt or restrict communicative exchanges without compelling reasons.[60] The right to have access to pertinent, public, information is also part of the core.[61]

It should be recalled that not every kind of speech is protected under Art 10. Hate speech,[62] incitement to violence,[63] and Holocaust denial are speech acts;[64] however, they fall outside the protection established by Art 10. For instance, in *Leroy v. France*,[65] the Court considered that a satirical cartoon representing the attack to the World Trade Center coupled with the caption "We have all dreamt of it ... Hamas did it!" was not worthy of protection, for it was merely a plea for terrorism. Holocaust denial, incitement to violence, and hate-speech cannot be the object of balancing, for they fall outside the scope of Art 10, even when cloaked as (pseudo)-artistic products or activities.

4. Fake News: A New Challenge

The ECtHR is now facing a new challenge: the dissemination of fake news and pervasive use of hate speech on social networks.[66] As the Court explains in

58. Flauss, *supra* note 8 at 815.
59. Jeremy Waldron, "Rights in Conflict" (1989) 99(3) Ethics 503.
60. *Lingens*, *supra* note 1 at para 41; *Şener*, *supra* note 1; *Markt intern Verlag GmbH and Klaus Beermann v Germany* (1989), 12 EHRR 161.
61. *Österreichische Vereinigung zur Erhaltung, Stärkung und Schaffung v Austria*, No 39534/07 (28 November 2013) at para 41.
62. See, e.g., *Norwood v the United Kingdom*, No 23131/03, [2004] XII ECHR 3. For an overview see Antoine Buyse, "Dangerous Expressions: the ECHR, Violence and Free Speech" (2014) 2 Intl & Comp L Q 491.
63. *Sürek v Turkey*, Nos 23927/94 and 24277/94, [1999] IV ECHR 353 at para 40.
64. The content is the thought expressed by the utterance (e.g. the *thought* that the Holocaust did not take place); the illocutionary force is the "use" of this content (e.g. the *assertion* that the Holocaust did not take place); the perlocutionary effects are the reactions caused by the utterance (e.g. the audience's dignity is offended). See Marina Sbisà, "Locution, Illocution, Perlocution" in Marina Sbisà & Ken Turner, eds, *Pragmatics of Speech Actions* (Mouton de Gruyter, 2013) at 25.
65. No 36109/03, (2 October 2008).
66. Danielle Keats Citron, *Hate Crimes in Cyberspace* (Harvard University Press, 2014); Alisdair Gillespie, "Hate and Harm: The Law on Hate Speech" in Andrei Savin & Jan Trzaskowski, eds, *Research Handbook on EU Internet Law* (Edward Elgar, 2015) at 488; James J Magee, *Freedom of Expression* (Greenwood, 2002) at 275-84; Alexander Tsesis, "Hate in Cyberspace: Regulating Hate Speech On the Internet" (2001) 38 SDLR 818; Jeremy Waldron, *The Harm in Hate Speech* (Harvard University Press, Cambridge, 2012); Jack M Balkin, "How Rights Change: Freedom of Speech in the Digital Era" (2004) 26(1) Sydney

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

Shtekel v. Ukraine, the Internet can be both an instrument for increasing freedom of expression and a danger for the protection of the other human rights:

> "The electronic network, serving billions of users worldwide, is not and potentially will never be subject to the same regulations and control. The risk of harm posed by content and communications on the Internet to the exercise and enjoyment of human rights and freedoms, particularly the right to respect for private life, is certainly higher than that posed by the press. Therefore, the policies governing reproduction of material from the printed media and the Internet may differ. The latter undeniably have to be adjusted according to the technology's specific features in order to secure the protection and promotion of the rights and freedoms concerned."[67]

Misinformation and propaganda are as old as time, but the impact that fake news had on central juridico-political events such as the U.S. Presidential elections and Brexit has raised the level of public awareness on this issue.[68] Most of our knowledge does not come from direct observation of facts because we tend to rely on reports and testimony. Therefore, the actual structure of the web represents a real risk: bad quality information overwhelms the global audience. Social networks—the "modern public squares"[69]—are often misunderstood as a reliable source of knowledge, even when they express unjustified or false beliefs.[70] The dissemination of fake news has two main reasons: the advertising revenue connected to

L Rev 5. See also *Report of the Special Rapporteur to the General Assembly on the Contemporary Challenges to Freedom of Expression*, OHCHR, 2016, A/71/373.

67. *Shtekel*, *supra* note 38 at para 63. This does not mean that, when a legal issue involves freedom of expression on the Internet, the competing rights always take the priority because no right shall be considered as exempt from balancing. The Court underscored this point in *KU v Finland*, No 2872/02, [2008] V ECHR 2872. See also EC, *Tackling online disinformation: A European Approach*, (2018) COM/2018/236 at 2.2 [*Tackling online disinformation*]: "some platforms have taken on functions traditionally associated with media outlets, entering the news business as content aggregators and distributors without necessarily taking on the editorial frameworks and capabilities of such outlets."

68. Richard Gunther, Paul A Beck & Richard Nielsen, "Fake News May Have Contributed to Trump's 2016 Victory" (2018) [unpublished, archived at: assets.documentcloud.org [perma.cc/WVT5-65YR]; DJ Flynn, Brendan Nyhan & Jason Reifler, "The Nature and Origins of Misconceptions: Understanding False and Unsupported Beliefs about Politics" (2017) 38(1) Political Psychology 127; DMJ Lazer et al, "The Science of Fake News" (2018) 359(6380) Science 5; Hunt Allcott & Matthew Gentzko, "Social Media and Fake News in the 2016 Election" (2017) 31(2) J Econ Perspectives 211; Andrew Guess, Brendan Nyhan & Jason Reifler, "Selective Exposure to Misinformation: Evidence from the Consumption of Fake News during the 2016 US Presidential Campaign" (2018) working paper, available at: www.dartmouth.edu [perma.cc/FL4L-4ZT6]; Aaron Quinn, "Fake News, False Beliefs, and the Need for Truth in Journalism" (2017) 31(1) Intl J Appl Phil 21; Sharyl Attkisson, *The Smear: How Shady Political Operatives and Fake News Control What You See, What You Think, and How You Vote* (Harper Collins, 2017) at 1ff; Sionaidh Douglas-Scott & Nicholas Hatzis, eds, *Research Handbook on EU Law and Human Rights* (Edward Elgar, 2017).

69. *Packingham v North Carolina*, 137 S Ct 1733 (2017).

70. Cass Sunstein, *On Rumors: How Falsehoods Spread; Why We Believe Them, What Can Be Done* (Farrar, Straus and Giroux, 2009).

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

the number of people accessing the information and the ideological and political strategy of disinformation.[71] The dangers of fake news are not limited to libel and hate speech:[72] *information overload* can disincentivize "quality journalism" and, paradoxically, this phenomenon might end up by frustrating the right of the citizens to be adequately informed.[73] By correctly informing the audience, news media ensure epistemic coverage to individuals. In a democratic republic, news media must provide a net of reliable beliefs to the audience.[74]

In the next few years, the ECtHR will play a fundamental role in supervising the filtering and take-down practices prescribed by the Contracting States, hopefully preventing gross violations of Art 10 by striking down illegitimate measures.[75] To be sure, one might contend that fake news poses a pseudo-problem. The Internet is the new "marketplace of ideas" and, therefore, the Court should not impose restrictions on freedom of expression. However, this argument *à la* Holmes is affected by a major flaw.[76] The ECtHR acts in a legal order which is already regulated by the Contracting States, in order to check, whether the national entities fulfill their obligations. Furthermore, the media market of platform providers resembles an oligopoly more than a highly competitive framework, despite the measures aimed at weakening of the barriers to entry.[77]

As the recent comparative inquiry of the Council of Europe shows, most European States have promulgated targeted regulations prescribing the removal from webpages and social networks of illegal content concerning child abuse,

---

71. Allcott & Gentzkow, *supra* note 68 at 214; Paul Bernal, "Fakebook: Why Facebook Makes the Fake News Problem Inevitable" (2018) 69(4) N Ir Leg Q 513; Jonathan M Smith, "Disinformation: A Taxonomy" (2011) IEEE Security and Privacy Magazine 1. The ECHR has approached this issue in *Centro Europa 7 Srl and Di Stefano v Italy*, No 38433/09 [2012] III ECHR 261; *Bowman v The United Kingdom* (1998), Eur Ct HR 175 at para 1.

72. See *Høiness v Norway*, No 43624/14, (19 March 2019) on the liability of Internet portals for allegedly offensive content; *Savva Terentyev v Russia*, No 10692/09, (4 February 2019) where the ECHR grants high-level of protection to insulting speech about the police published on a blog; *Egill Einarsson v Iceland (No 2)*, No 31221/15, (17 July 2018), on defamatory content posted on a Facebook page during a police investigation.

73. *Tackling online disinformation*, *supra* note 67 at 1: "Our open democratic societies depend on public debates that allow well-informed citizens to express their will through free and fair political processes. Media have traditionally played a key role in holding public authorities to account and in providing the information that enables citizens to form their own views on societal issues and actively and effectively participate in democratic society. [. . .] Disinformation erodes trust in institutions and in digital and traditional media, and harms our democracies by hampering the ability of citizens to take informed decisions." See also Commissioner for Human Rights, *Human Rights in a Changing Media Landscape* (Council of Europe Publishing, 2011); Andrea Renda, "The legal framework to address "fake news": possible policy actions at the EU level" (2018), online (pdf): *European Parliament* europarl.europa.eu [https://perma.cc/WX38-L525].

74. Axel Gelfert, "Fake News: A Definition" (2018) 38(1) Informal Logic 84 at 88.

75. Council of Europe, *Comparative Study on Blocking, Filtering and Take-Down of Illegal Internet Content* (January 2017).

76. *Abrams v United States*, 250 US 616, 630 (1919) (Holmes J dissenting): "[W]hen men have realized that time has upset many fighting faiths, they may come to believe even more than they believe the foundations of their own conduct that the ultimate good desired is better reached by free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market."

77. Daniel Mandrescu, "Applying (EU) Competition Law to Online Platforms: Reflections on the Definition of the Relevant Market(s)" (2018) 41(3) World Competition 453.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

terrorism, hate crimes, intellectual property, and copyright. The primary trend in Europe is to impose a procedure of 'notice and take-down' whereby the contentious content shall be removed by the host after publication in response to allegations of illegality. However, most legal systems also require Internet hosts, social networks, and on-line service providers to act *ex ante*, by filtering in the phase between upload and publication.[78]

The legal issues related to these practices are incredibly delicate. Sharing a content on the Internet is always the result of a cooperation between several actors: a) the agent who posts the content, b) the host who provides a domain-based platform (e.g. Facebook, Twitter, Word-Press), c) the audience, and (d) whoever facilitates or checks this process (e.g. ISP providers).[79] Web-expression is radically different from ordinary communicative exchanges.[80] The audience plays an active role in the uptake of the content, and in responding with likes and further comments. In principle, both the agent who writes a post or comment and the host are accountable, although, according to the European Commerce Directive and the Audio-visual Media Services Directive, web operators benefit of *limited* accountability.[81]

Many hosts and intermediaries *act as editors*, making a profit out of what users share on their platforms, but, strictly speaking, they lack the legal status of a professional editor. In this case, the liabilities are not clearly and consistently established by the Member States. On that note, in the recent case of *Rebechenko v. Russia*,[82] the ECtHR proclaimed that bloggers, too, exercise the function of "public watchdogs." However, any enthusiasm for the analogies between journalists and bloggers as well as editors and social networks should be tempered. In Kate Klonick's words, "a central piece is missing in the comparison to an editorial desk: platforms do not actively solicit specific types of content, unlike how an editorial desk might solicit reporting or journalistic coverage. Instead, users use the site to post or share content independently."[83]

Grave concerns about the risk of censorship in disguise and "collateral overcensorship" have already emerged in Europe, where the Member States have started to issue special regulations against fake news.[84] Germany has adopted

78. Kate Klonick, *supra* note 30 at 1635; James Grimmelmann, "The Virtues of Moderation" (2015) 17 Yale JL & Tech 42 at 63-70.

79. Seth C Lewis & Oscar Westlund, "Actors, Actants, Audiences, and Activities in Cross-media News Work" (2015) 3(1) Digital Journalism 19.

80. See *Report of the Special Rapporteur*, *supra* note 66 at para 4 (on the role of intermediaries). See also Rasmus K Nielsen, "News Media, Search Engines, and Social Networking Sites as Varieties of Online Gatekeepers" in Chris Peters & Marcel Broersma, eds, *Rethinking Journalism Again* (Routledge, 2016) at 81; Rasmus K Nielsen & Sarah A Ganter, "Dealing with Digital Intermediaries: A Case Study of the Relations between Publishers and Platforms" (2018) 20(4) New Media & Society 1600.

81. Lorena M Woods, "Video-sharing Platforms in the revised Audiovisual Media Services Directive" (2018) 23(3) Communications L 127; Etienne Montero & Quentin Van Enis, "Enabling freedom of expression in light of filtering measures imposed on Internet intermediaries: Squaring the circle?" (2011) 27 Computer L & Security Rev 21.

82. No 10257/17, (16 April 2019).

83. Klonick, *supra* note 30 at 1660.

84. For an overview see Tarlach McGonagle, "'Fake news': False Fears or Real Concerns?" (2017) 35(4) NQHR, 203-09. On "collateral censorship" see Klonick, *supra* note 30 at 1608:

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

the Network Enforcement Act in 2017 which requires social media and platform providers to swiftly remove fake news, hate-inciting content, and any "obviously illegal" content (an excessively vague and ambiguous term indeed), at the latest within twenty-four hours of being notified, imposing fines up to 50 million Euros.[85]

In France, the so-called "law against the manipulation of information", which introduced three new articles (L. 112, L. 163-1 et L. 163-2) in the Electoral Code (*Code Électoral*), was challenged in the national courts. These legal provisions prohibit the dissemination of fake news about the political process during the three months preceding elections and prescribe sanctions for providers publishing false information that might influence the voter's preferences. It comes as no surprise that the *Conseil Constitutionnel* upheld the French network enforcement act, considering that the measure was not disproportionate since it was limited to the elections and concerned only information of public interest.[86]

The European Union, too, has begun to wage war against fake news, by promoting the EU Code of Practice on Online Disinformation in 2018. The EU Code, on the one hand, proposes AI aided systems for preventing disinformation and for tracking and removing automatically illegal content from online platforms. On the other hand, the code introduces a number of measures—such as the external review of an independent third party—aimed at preventing potential side effects such as the inhibition of *bona fide* research and the possibility of political manipulation.

On 26th April 2018, the European Commission adopted a Communication on "tackling online disinformation"[87] which proposes guidelines for preventing fake news and disinformation through an intricated system of strategic measures: a) the use of new technologies, such as blockchains and AI systems, for tracking the origin of the information; b) the introduction of an independent European network for checking the quality of the information; c) several measures aimed at fostering education and "media literacy";[88] d) cyber-security safeguards for the elections; and (e) a plan of instructing journalists and public media on how they should act in a digital environment.

There is an actual risk that, under specific interpretations, the abovementioned legislative interventions might offer a pretext to national governments and EU institutions for over-censorship and ideological manipulation. For instance, improving the "media literacy" of journalists, on the paper, might sound like a splendid idea, but it could quickly degenerate into forms of libertarian paternalism aimed at brainwashing the fourth estate. True enough, the risk that fake news

"Collateral censorship occurs when one private party, like Facebook, has the power to control speech by another private party, like a Facebook user."

85. *Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken* (Netzwerkdurchsetzungsgesetz—NetzDG), Bundesgesetzblatt, Jahrgang 2017, Teil I, Nr 61.

86. Conseil Constitutionnel, Décision n° 2018-773 DC du 20 décembre 2018.

87. *Tackling online disinformation*, *supra* note 67. The European Commission announced the plan of forming high-level expert groups working on fake news as early as November 13, 2017.

88. See Council of Europe, "Developing Competences for Democratic Culture in the Digital Era" (2017), online (pdf): www.coe.int [perma.cc/DX3H-XGBG].

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

might harm other fundamental liberties, such as privacy, is considerable, especially considering the powerful effect of multi-media contents, the degree of anonymity of the users,[89] and the persistence of the information on the web.[90] However, the battle against fake news could result in disproportional censorship and political interference if the notion of fake news is defined by legislators in vague and broad terms such as 'unreliable information that can represent a threat to life, public order, public safety, health, and vital elements of social and transportation structure', as the Russian law notably does.[91] If the definition of "fake news" is too broad, vague, and ambiguous then, in principle, any content that does not meet with the governmental approval could be legitimately removed, and the system of sanctions would become unpredictable.

It behooves the ECtHR, as an independent judiciary, to have the last words on these fundamental issues concerning human rights; however, we might wonder whether the traditional approach based on categorization and sliding-scales is still suitable for dealing with web-expression and fake news. Have we already reached such radical change in our ways of expressing ourselves, as "digital speakers" that requires the abandonment of the more traditional categories for balancing Art. 10? Is it still possible to rely on "fixed" stipulations and categories? Wouldn't it be better to decide intuitively, case-by-case, without the ambition of drawing general lines? Should "fake news" be considered as an autonomous category of expression, and at what level of the scale should it be placed? The current analysis will propose arguments in favor of the traditional categorization-based approach. I will suggest that a minimal definition of "fake news" acts as a shield against excesses of the censorship machine. Furthermore, several subspecies of fake-news and several standards for balancing web-expression will be disentangled.

There are good reasons to think that the Court might still rely on categorization, for this seems to be the solution preferred also by EU Member States. National regulations tend to split web-expression into various sub-categories (e.g. child pornography, illegal online gambling, intellectual property, and personal data protection)[92] ascribing different levels of legal protection to each of them. Hence, the ECtHR might decide to associate systematically every sub-category of "web-expression" and "fake news" with specific balancing tests. The most

---

89. *Shtekel, supra* note 38 at para 63; *Delfi v Estonia*, No 64569/09, [2015] II ECHR 586 at para 132 [*Delfi*]; *Ahmet Yldirim v Turkey*, No 3111/10, [2012] VI ECHR 2012 at para 48 [*Ahmet Yldirim*]. See also Mireille Hildebrandt, "Primitives of Legal Protection in the Era of Data-Driven Platforms" (2018) 2(2) Geo L Tech Rev 2(2) at 255: "Due to their distributed, networked, and data-driven architecture, platforms enable the construction of invasive, over-complete, statistically inferred, profiles of individuals (exposure), the spreading of fake content and fake accounts, the intervention of botfarms and malware as well as persistent AB testing, targeted advertising, and automated, targeted recycling of fake content (manipulation)."

90. See also, *Google Spain SL, Google Inc v Agencia Española de Protección de Datos, Mario Costeja González*, C-131/12, [2014] ECR 317.

91. Federal Laws of 18.03.2019 Nos 31-FZ and 30-FZ On Amendments to Article 15-3 of the Federal Law On Information, Information Technologies and on Information Protection.

92. See, e.g., David Erdos, "European Data Protection Regulation and Online New Media: Mind the Enforcement Gap" (2016) 43(4) IJLS 534 at 538ff; Brandon Hanlon, "The 'European Approach' to Fighting Disinformation: Lessons for the United States" (2018), online (pdf): *Alliance for Securing Democracy* securingdemocracy.gmfus.org [perma.cc/NK2V-NHUK].

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

demanding tests and, accordingly, the highest level of protection (strict scrutiny) might be granted to online professional newspapers, whereas a lower level of protection and more lenient tests (e.g. intermediate scrutiny) might be associated to purely commercial speech. Then, with systematic effort, every category of "web-expression" might be placed at different levels of the scale. There are at least three recent decisions of the ECtHR that move in this direction.

In *Pirate Bay*, the Court upheld a Swedish ban on specific peer-to-peer downloading systems—viz sharing "torrent" files—that violated copyright protection according to the intermediate scrutiny. The ban was deemed consistent with the legitimate aim pursued by the legislator.[93] The intermediate scrutiny was "activated" by the specific category of 'informational piracy,' which, under its manifestly illegal nature, precludes heightened levels of scrutiny. Furthermore, the presence of an economic interest in spreading digital contents played a fundamental role in "tuning" the Court's assessment toward the lowest levels of the scale.

The ECtHR has resorted on the traditional approach (namely categorical balancing) also in the recent case of *Cengiz and Others v. Turkey*.[94] By applying a rational basis with bite test, the Court expressly held that the Turkish government's restrictions on *YouTube* broadcasting were aimed at affecting the citizens' right to receive information. First, the restrictive measure was not grounded on an enacted statute. What is more, the means/end rationality requirement was afflicted with *over-inclusiveness*.[95] National authorities are not allowed to block an entire website if only the content of a few hosted pages is illegal therefore, the restrictive measure was too broad compared to its underlying reason. The Court also considered that, strictly speaking, the case at hand was not an instance of purely commercial speech. In that specific context, *YouTube* was acting as a public intermediary and the disputed content concerned the political figure of General Atatürk, the first President of Turkey (1923-38). Therefore, a heightened burden of justification was appropriate.

Finally, in the *case of Magyar Jeti Zrt v. Hungary*[96] the ECtHR considered that the *dissemination* of defamatory content cannot be equated to the *mere posting* of a hyperlink that opens a defamatory content; in cases involving hyperlinks, any restrictive measure shall satisfy the three-pronged necessity test.[97] Furthermore, the individual assessment on a hyperlink case must ponder eight relevant factors: "i) did the journalist endorse the impugned content; ii) did the journalist repeat the impugned content (without endorsing it); iii) did the journalist merely include a hyperlink to the impugned content (without endorsing or repeating it); iv) did the journalist know or could the journalist reasonably have

93. *Fredrik Neij and Peter Sunde Kolmisoppi (The Pirate Bay) v Sweden*, No 40397/12 (19 February 2013).

94. Nos 48226/10 and 14027/11, [2015] VIII ECHR 147 at para 63-7 [*Cengiz*].

95. See Frederick Schauer, *Playing By the Rules: A Philosophical Examination of Rule-Based Decision Making in Law and in Life* (Clarendon Press, 1991) at 31ff.

96. No 11257/16 (4 December 2018) at para 56-85 [*Magyar*].

97. For a critical stance on the necessity test, see Janneke Gerards, "How to Improve the Necessity Test of the European Court of Human Rights?" (2013) 11 Intl J Const L 466.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

known that the impugned content was defamatory or otherwise unlawful; v) did the journalist act in good faith, respect the ethics of journalism and perform the due diligence expected in responsible journalism?"[98]

Now, categorial balancing would be a promising path for dealing with the expression on the web and fake news only if integrated with a comprehensive theoretical development that identifies new sub-categories of web-expression and a *minimal definition* of the concept "fake news", a philosophical puzzle that the ECtHR is called to resolve.

It is quite challenging to stipulate a foundational, precise, and neutral working definition of the compound term "fake news".[99] However, such a definition is necessary for both preventing the risk of censorship in disguise, and for a reasoned elaboration of the ECHR principles. Without a neutral definition of fake news, any balancing operation on this matter would be like drawing lots,[100] and the crusade against the spread of bad quality information on the web might gradually lead to abuses. "Fake news" cannot become a "catch-all term"[101] which empowers biased censors to act as the ultimate arbiters of truth.[102]

As plausibly suggested by the philosophers Romy Jaster and David Lanius, the best solution for a stipulative approach is to start with a *minimal* two-fold notion of "news" as "newly received or noteworthy information, especially about recent events" and, in a more specific usage, as a "broadcast or published report of news."[103] Then, it could be added that the qualification "fake" means that the news is alternatively: a) false; b) true but highly misleading; c) made without any intention of seeking the truth. The presence of the third element c) is fundamental for it captures the idea that fake news is "misleading *by design*".[104] A piece of fake news is intended to have a broad, increasing circulation, deceive a targeted audience, and to exploit prejudice and biases of the audience (e.g. biases about some social issues such as the impact of immigration) for reaching lucrative or political goals (e.g. influencing an election).[105]

This definition has the virtue of separating *fake news* from *editorial errors gone viral*. Fake news comprises either *outright false reports* (e.g. the news that Obama "will refuse to leave his office if Trump is elected" shared by

98. *Magyar*, *supra* note 96 at para 77.
99. Nabiha Syed, "Real Talk About Fake News: Towards a Better Theory for Platform Governance" (2017) The Yale LJ 337; Claire Wardle, "Fake News. It's Complicated" (2017), working Paper available at: medium.com [perma.cc/FFE3-WKVQ].
100. See Gelfert, *supra* note 74 at 85: "if one holds that the term refers to a distinct and identifiable phenomenon, it is all the more important to attempt to come to conceptual grips with it and, if necessary, put forward a definition that can then be refined in the course of future scholarly debate."
101. McGonagle, *supra* note 84 at 203-09.
102. *Tackling online disinformation*, *supra* note 67 at para 3.1.2 alludes at the opportunity of reliable "independent fact-checkers", whose task is to analyze the credibility of the content, as well as the correctness of the creation and dissemination process.
103. Romy Jaster & David Lanius, "What is Fake News?" (2018) 127(2) Versus 207.
104. See Gelfert, *supra* note 74 at 86, 109-12.
105. Syed, *supra* note 99 at 352.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

Burrardstreetjournal.com on Sept 7, 2016)[106] or *misleading information* (for instance, the report that, after Hurricane Dorian, *bricks* of cocaine washed up on Florida's beaches was misleading because only *one single brick* was found).[107] In both cases, the "fake" report is aimed at *deceiving the audience* with an *utter disregard of the truth*.[108] In this respect, fake news is radically different from journalistic errors gone viral. Viral journalistic errors, although false, concern issues of general interest and have been obtained through *bona fide* accurate research.[109] The journalist genuinely believes in the truth of what she says (i.e. the speaker is truthful, although the content is false), and her conclusion is supported by evidence.[110] Genuine journalistic errors gone viral deserve special protection, for excessive constraints imposed on professional journalists, acting *bona fide*, might frustrate their fundamental role as watchdogs of the democracy,[111] as well as their ability to *debunk* fake news. The *Stoll v. Switzerland*[112] case is an indication that the ECtHR is perfectly aware of this risk.

According to the minimal definition, not every piece of fake news is straightforwardly *false* and, conversely, not every false report is "fake news". Sensational

---

106. "President Obama Confirms He Will Refuse To Leave Office If Trump Is Elected", *Burrard Street Journal* (7 September 2016), online: burrardstreetjournal.com [perma.cc/RW3S-2K5W].

107. Minyvonne Burke, "Bricks of cocaine wash up on Florida beaches from Hurricane Dorian waves", *NBC News* (4 September 2019), online: nbcnews.com [perma.cc/5GMW-AWSU].

108. Fake news can be regarded, to some extent, as lies; lies, in turn, can be considered as moral wrong when they are an intentional source of deception. See Seana V Shiffrin, *Speech Matters: On Lying, Morality, and the Law* (Princeton University Press, 2014) at 12; Miranda Fricker, *Epistemic Injustice* (Oxford University Press, 2007) Chapter VI. The definition of "fake news" proposed by Jaster, *supra* note 103 at 207 might count as an analytical refinement of the one proposed by the European Commission. See *Tackling online disinformation*, *supra* note 67 at para 2.1: "Disinformation is understood as verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public, and may cause public harm. Public harm comprises threats to democratic political and policy-making processes as well as public goods such as the protection of EU citizens' health, the environment or security. Disinformation does not include reporting errors, satire and parody, or clearly identified partisan news and commentary."

109. *Goodwin v UK* (1996), 22 EHRR 123 at para 39.

110. *Times Newspapers Ltd v the United Kingdom* (nos 1 and 2), Nos 3002/03 and 23676/03, [2009] I ECHR 377 at para 27: "In the light of its accessibility and its capacity to store and communicate vast amounts of information, the Internet plays an important role in enhancing the public's access to news and facilitating the dissemination of information in general. The maintenance of Internet archives is a critical aspect of this role and the Court therefore considers that such archives fall within the ambit of the protection afforded by Article 10."

111. The ECHR emphasized on many occasions that "responsible journalism" is highly protected by the Convention. See, e.g., *Flux and Samson v Moldova*, No 28700/03 (23 October 2007); *Standard Verlags GmbH v Austria*, No 34702/07 (10 January 2012). Similar principles are endorsed by Council of Europe, PA, Res 2143, *Online Media and Journalism: Challenges and Accountability* (2017). See also Dirk Voorhoof & Hanne Cannie, "Freedom of Expression and Information in a Democratic Society: The Added but Fragile Value of the European Convention on Human Rights" 72(4) (2010) ICG at 417.

112. [2007] V ECHR 207 at 103-04: "These considerations play a particularly important role nowadays, given the influence wielded by the media in contemporary society: not only do they inform, they can also suggest by the way in which they present the information how it is to be assessed. In a world in which the individual is confronted with vast quantities of information circulated via traditional and electronic media and involving an ever-growing number of players, monitoring compliance with journalistic ethics takes on added importance."

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

and purely fictional stories, albeit false, do not count as fake news if they are presented as unreal contents, shared as a pure source of entertainment.[113] Websites that deliberately and unabashedly offer purely fictional stories are not spreading, strictly speaking, fake news. With an eye to the responsibility of the platform providers and social networks, the flagging strategies employed deserve close attention. These intermediaries might flag inaccurate content either as "disputed" or as "false," and the choice between these two options has a substantial impact on the audience.[114]

In this connection, *satire* deserves close attention. Since the times of the Greek comic playwriter Aristophanes, satirical expression plays a fundamental role in encouraging the political debate. This historical evidence shows that purely fictional political content might become harmful only when it is misleadingly used to fool the audience. The fictional character of the content shall be clearly stated in the text. For instance, this type of warning was intentionally omitted when the webpage *wtoenews5.com* reported the fake news that Pope Benedict supported Donald Trump for the presidential elections.[115] By contrast, the news that Rep. Ilhan Omar walked out of a 9/11 Memorial Service, published by *Bustatroll.org*, was unequivocally portrayed as a piece of satire by the editors.[116] Ultimately, under the proposed account, genuine satire falls outside the scope of the term "fake news".

The minimal definition of "fake news" includes *misleading* statements which, while being *literally* true, can be associated with implicit false content and, therefore, misrepresent some facts (what we might dub "distorted news").[117] Misleading statements are spurious, although not, strictly speaking, counterfeit. An instance of a misleading statement is the use of hyper-partisan interpretations of facts or the malicious selection of facts. These kinds of statements might affect significantly the right to be correctly informed.

Lastly, according to the proposed definition, the notion of fake news includes those statements made without any concern for the truth, where the content is made up by the speaker only to shape the beliefs of her audience. Let us imagine, for instance, a journalist that, in an acclaimed online newspaper, describes NGO's vessels involved in the operation of rescuing immigrants at the high seas as "sea-taxis." This misleading description would be fabricated without any intention to

113. Matthew RX Dentith, "The Problem of Fake News" (2017) 8(1/2) Pub Reason 65-69 at 66. See also Dentith, *In Defence of Conspiracy Theories* (PhD Dissertation, University of Auckland, 2012) [unpublished].

114. Gordon Pennycock & David G Rand, "The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Stories Increases Perceived Accuracy of Stories Without Warnings" (2017) Working Paper, online (pdf): osf.io [perma.cc/5BEG-T9LJ].

115. Allcott & Gentzkow, *supra* note 68 at 214.

116. Colon Crusher, "Ilhan Omar Walks Out Of 9/11 Memorial Service", *Bustatroll*, online: bustatroll.org [perma.cc/753S-U2E7].

117. Lance Bennet & David Livingston, "The Disinformation Order: Disruptive Communication and the Decline of Democratic Institutions" (2018) 33(2) Eur J Comm 122; Yochai Benkler, Robert Faris & Hal Roberts, *Network Propaganda: Manipulation, Disinformation and Radicalisation in American Politics* (Oxford University Press, 2018) at 6.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

seek the truth, for the sole purpose of eliciting a negative bias towards NGO's rescue operations and, plausibly, also towards immigration in general.

To close the square, we can say that fake news, as a broad category, includes lies (i.e. false statements made with the intention of deceiving), true but (intentionally) misleading statements (recall the "sea-taxis"), and statements made without any concern for the truth. The glue that keeps these notions together is the lack of truthfulness of the speaker.

This minimal definition will pay good dividends in several ways. On the one hand, it "saves" earnest and accidental mistakes but, on the other, it does not restrict the notion of fake news to "completely fabricated stories" and "contents that are 100% false",[118] a stipulation that seems too narrow. Furthermore, it maps widespread intuitions about fake news, such as the presence of malicious intent, the lack of accuracy, the absence of factual basis, and the idea that a piece of fake news is purported as true. In this way, the minimal definition rules out several categories that are not harmful, such as satire, (subjective) value-judgments, and urban legends. In fact, value-judgments and personal opinions—which, according to the jurisprudence of ECtHR are protected by Art 10—fall outside the minimal definition of "fake news" advocated by the current analysis. Why should we consider the subjective opinions or value-judgments of a blogger as fake news, if they are not purported as a truth (or truthful) description of facts? Finally, the minimal definition here proposed is neutral since it is not charged with moral and political notions.

Now, a minimal definition of "fake news" might be a starting point for preventing over-censorship, but it is not sufficient. As mentioned above, the categorial balancing and the sliding-scales approach require a more fine-grained characterization through precise sub-categories that provide a "scalar framework" for the Court's balancing operations.[119] The peculiarities of the Internet as a virtual space require an altogether *new* sub-categorization. A perspicuous analysis of fake news should be informed by the mediating digital and cultural infrastructure that facilitates the online circulation.[120]

First, it can be useful to distinguish different types of websites, based on their primary contents—arts, business, health, recreation, and politics, to name just a few—because it seems reasonable that the standards for assessing bad quality information and its potential harmful effects vary according to this parameter.[121] The structure of the multi-media service shall also be taken into account. On this point, the Visual Media Service Directive proposes a legal regime that distinguishes between "linear services"—namely, scheduled audio-visual services

118. Gelfert, *supra* note 74 at 96.
119. See Jacob Rowbottom, "To Rant, Vent and Converse: Protecting Low Level Digital Speech" (2012) 71(2) Cambridge LJ 355 at 376.
120. Philip E Agre, "Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI" in Geof Bowker, Les Gasser, Leigh Star & Bill Turner, eds, *Bridging the Great Divide: Social Science, Technical Systems, and Cooperative Work* (Erlbaum, 1997).
121. See Hunt Allcott, Matthew Gentzkow & Chan Yu, "Trends in the Diffusion of Misinformation of the Social Media" (2019) R&P at 4.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

(e.g. BBC) or on-demand services (e.g. Netflix)—and "user-generated social media" (e.g. Twitter).[122] Only the former is required to assume editorial responsibility. The ECtHR might wish to endorse these classifications. More broadly, considerations about the corporate identity (namely, the values endorsed by the company) may figure in the Court's reasoning.

Second, balancing operations should look attentively at the importance of the website and separate highly ranked websites from small websites.[123] Generally speaking, a piece of fake news spread by a major website has a stronger impact than a piece of fake news divulgated by a small webpage, although the end-to-end infrastructure of the Internet and the audience tendency of submitting to purported authorities might defeat this generalization under specific circumstances. More broadly, the impact factor of a piece of fake news (which depends on the aggregated quantity of likes, shares, and comments) shall be measured. News coming from top-ranked journals (The Sunday Times, BBC, Daily Mirror, Le Monde) and fact-checking websites (such as "FakeCheck.org", "Snopes.com", and "mercola.com") shall be treated separately, both for their positive debunking capacity and for the detrimental potential drawbacks of their possible mistakes. Top-ranked journals, in particular, are often considered as authoritative sources by the audience and, therefore, granted higher credibility.

Finally, the peer-to-peer distribution and, in general, the reaction of the audience should be assessed through consequential reasoning. If we consider the peer-to-peer distribution, it is hard to see how false information, tweets, and comments shared by troll farms, troll brigades, and bots might count as protected expression.[124] "Troll farms" are organized teams of users paid to spread fake news and inflammatory, offensive, provocative, or senseless content on social media and platforms, in order to create conflict and disruption. "Troll brigades", by contrast, are State-sponsored anonymous Internet political commentators that rely on sock-puppets, large-scale orchestrated disinformation, propaganda campaigns, and hate speech, to exercise influence on the political and social life of other states.[125] "Bots" are automated advertising engines that commonly spread spam, sensational content, and inflammatory speech (e.g. the news that the Pope supports Trump). As the Court clarified in *Cengiz and Others v. Turkey*, Art 10 ECHR protects the *user-generated expressive activity*,[126] and not *bot- or troll-generated expressed activity.*

To be sure, the ECtHR has grappled with web-expression in a few other recent cases, including the landmark decisions *Delfi v. Estonia* and *MTE Index v.*

---

122. EC, *Council Directive 2007/65/EC on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the pursuit of television broadcasting activities*, [2007] OJ L 332/27.

123. *Ibid* at 4.

124. Clint Watts, *Messing with the Enemy. Surviving in a Social Media World of Hackers, Terrorists, Russians, and Fake News* (Harper, 2018) at 163ff.

125. Peter Pomerantsev & Michael Weiss, *The Menace of Unreality. How the Kremlin Weaponizes Information, Culture and Money. Institute of Modern Russia* (2014), online (pdf): http://www.interpretermag.com [perma.cc/BS62-9V6D].

126. *Cengiz*, *supra* note 94 at paras 49, 52.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

Hungary.[127] However, the Court's reasoning mainly addresses hate speech and defamatory contents, without meeting the challenge of fake news head-on. In *Delfi*, the Grand Chamber held with a striking majority (15-2) that companies are liable for the late removal of unlawful, offensive, vulgar, and violent content from their website. The Court emphasized the commercial nature of the portal and, correspondingly, the economic interest of the company in the posting of comments. *Delfi* was not acting merely as a disinterested service provider, for it allowed false, offensive, comments that were not protected by freedom of expression in order to gain economic benefits. In this way, the Court substantially equated the company's liability to the liability of a journalistic publisher, in consideration of the fact that the company was operating *de facto* as a discloser of information. The ECtHR applied the same rationale *MTE Index v. Hungary*, which, in this case, led to the opposite solution because the relevant facts were radically different. MTE was a non-profit service provider therefore, it was not economically interested in the number of comments. Furthermore, the negative comments that were published on the MTE's platform did not amount to hate speech.[128]

These are important and most welcome achievements. The presence of an economic interest is a marker for 'commercial expression', which attracts low levels of protection and less demanding tests. However, this solution cannot be hailed as a turning point in ECtHR jurisprudence. As adumbrated above, in future analogous cases the Court might wish to consider a more extensive set of variables directly related to 'fake news' in order to move beyond the decisional strategy based solely on the dichotomy commercial/noncommercial and the possible presence of hate-speech.[129] This seems quite trivial if we consider that the vast majority of on-line service providers *are* economically motivated entities. Furthermore, the social costs of bad quality information (less accurate and mistaken beliefs, skepticism toward news producers, and unreliable information regarding the political process) should be balanced against: i) the interest of the news consumers in enjoying the reading of fake news, as a sheer form of

---

127. *Delfi*, *supra* note 89; *MTE Index v Hungary*, No 22947/13 (2 February 2016). See also *Rolf Anders Daniel Pihl v Sweden*, No 74742/14 (9 March 2017) [*Pihl*] on defamatory blogs and *Ashby Donald and Others v France*, No 36769/08 (10 January 2013) at para 34 on the publication of photographs on a fashion website. See, generally, Bart van der Sloot, "Welcome to the Jungle: The Liability of Internet Intermediaries for Privacy Violations in Europe" (2015) 6(3) JIPITECL 211; Van der Sloot, "The Practical and Theoretical Problems with 'Balancing' Delfi, Coty and the Redundancy of the Human Rights Framework" (2016) 23(3) MJECL 439 at 448; Tatiana E Synodinou, "'Intermediaries' Liability for Online Copyright Infringement in the EU: Evolutions and Confusions" (2015) 31(1) CL&SR 57 at 63; Urs Gasser & Wolfgang Schulz, "Governance of Online Intermediaries Observations from a Series of National Case Studies" in *Governance of Online Intermediaries Observations from a Series of National Case Studies* (The Berkman Center for Internet & Society, 2015); Ronen Perry & Tal Zarsky, "Liability for Online Anonymous Speech: Comparative and Economic Analyses" (2014) 5(2) JETL 205.

128. Eileen Weinert, "MTE v Hungary: The first European Court of Human Rights Ruling on Liability for User Comments after Delfi AS v Estonia" (2016) 27(4) Entertainment LR 27 135.

129. *Pihl*, *supra* note 127 ascribes a pivotal role to the small and non-profit nature of the website as a reason for not imposing liability on Internet intermediaries.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

entertainment; ii) the interest in protecting the technical functioning and the access to the Internet, "regardless of frontiers";[130] and (iii) the freedom of expressing subjective value judgments.[131] The ability to enhance civic participation and the expression of political sentiments (iv) should be examined, too. It would suffice to recall the solidarity impact of the hashtag initiative "#jesuischarlie" in the aftermath of the terroristic attack to Charlie Hebdo.[132]

Algorithmic filtering systems endorsed by companies or prescribed by the Member States in particular deserve close attention. On the one hand, there is an actual risk that excessive sanctions and strict liability will create direct incentives for platforms to silence web-expression through a widespread removal of contents and a policy of over-acquiescence to national executives.[133] On the other, filters based on algorithms—which, in turn, are developed in accordance with the goal of profit maximization—tend to reinforce group "non-deliberative polarization" on the web (the so-called "echo-chambers") by both increasing exchanges of ideas among like-minded users (e.g. persons that share the same political views, or the same opinions on climate changes) and exposing users to contents that match their personal tastes.[134] What is more, there are good reasons to suspect that a purely "algorithmic" approach to filtering is inaccurate.[135] Algorithms are still scarcely able to deal with complex linguistic phenomena connected to the pragmatics of communication, such as irony, satire, humor, and absurdity.[136] They are still unable to compute the emotional valence of an expression efficiently and to detect how

---

130. *Ahmet Yldirim*, *supra* note 89 at para 67.

131. Neil Levy, "The Bad News About Fake News" (2017) 6(8) Social Ep Rev 20; Allcott & Gentzkow, *supra* note 68 at 214. See also *Tierbefreier e V v Germany*, No 45192/09 (16 January 2014) at para 56 on the online exchange of ideas is protected by Art 10 until it becomes a sheer personal attack.

132. Patrick Leerssen, "Cut Out By The Middle Man: The Free Speech Implications Of Social Network Blocking and Banning" (2015) 6 EU JIPITEC 100 at 101.

133. Klonick, *supra* note 30 at 1616ff, 1651, 1667.

134. Cass R Sunstein, *Republic.com 2.0* (Princeton University Press, 2007) at Chapter III; Syed, *supra* note 99 at 346f. On polarization *qua* socio-psychological phenomenon see Daniel J Isenberg, "Group Polarization: A Critical Review and Meta-Analysis" (1986) 50 J Personality & Soc Psychol 1141.

135. Niall J Conroy, Victoria L Rubin, & Yimin Chen, "Automatic Deception Detection: Methods for Finding Fake News" (2015) 52(1) Proceedings of the Association for Information Science & Technology 1; Song Feng, Ritwik Banerjee & Yejin Choi, "Syntactic Stylometry for Deception Detection" in ACL, *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2* (Jeju, 2012) at 171; Rada Mihalcea & Carlo Strapparava, "Making Computers Laugh: Investigations in Automatic Humor Recognition" in HLT, *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing* (Stroudsburg, 2005) at 531. See also Sridhar Ramaswamy (Senior Vice President, Ads & Commerce of Google), "How We Fought Bad Ads in 2015" (21 January 2016), available at: https://blog.google/technology/ads/better-ads-report/; Mark E Zuckerberg, "Building Global Community" (16 February 2017), online: www.facebook.com [perma.cc/NDW4-TZD4]; Lilian Edwards & Michael Veale, "Slave to the Algorithm? Why a 'Right to Explanation' is Probably Not the Remedy You are Looking for" (2017) 16 Duke L & Tech Rev 18.

136. Eileen Fitzpatrick, Joan Bachenko & Tommaso Fornaciari, "Automatic Detection of Verbal Deception" (2015) 8(3) Synthesis Lectures on Human Language Technologies 1; Byron C Wallace, "Computational Irony: A Survey and New Perspectives." (2015) 43(4) AI Rev 467.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

pragmatic markers yield implicit contents.[137] Eventually, politically biased algorithms can "disproportionally favor people with power over individual users" and shrink the right to equal access and participation.[138] For all these reasons, the structure of the filtering system shall also figure among the factors for balancing fake news and, more broadly, web-expression.

If the notion of intention plays a fundamental role in defining "fake news", as I have suggested, the ECtHR might enter deep water. the possibility of relying extensively, for the first time, on "intention-based tests", namely, complex judicial standards for evaluating the intentions undergirding expressed content in order to ascertain the presence of actual malice. What is more, the Court might also rely on an "epistemic entitlement test", which checks whether the content of the expression is epistemically justified in terms of transparency of the information, trustworthiness of the source, and soundness of the inferential patterns that led to its creation. This solution would really amount to a radical change in the ECtHR jurisprudence, for human rights courts generally hesitate to focus on the motives of the speakers, because motives are difficult to prove.[139]

Now, one question remains open: could we trace back the notion of fake news to the more general, old, and broader category, of false news? False news, propaganda, and other forms of weaponization of information are not at all new. One might recall Octavian's propaganda campaign against Antony based on the dissemination of slogans imprinted on coins (1st Century BC), or the *War of the Worlds* radio drama (1938), narrated by Orson Wells, which tricked the listeners into believing that a Martian invasion was actually taking place.[140]

The notion of *false* news is not unprecedented in the jurisprudence of the ECtHR. For instance, in her insightful essay, Irina Katsirea mentions *Raëlien Suisse* among the prominent cases of non-digital dissemination of fake news. In that case, the ECtHR was dealing with a poster campaign aimed at convincing people about the existence of Aliens.[141] However, there are several properties that enable *distinguishing* between fake news and false news *tout court*: a) the fact that "new technologies can be used, notably through social media, to disseminate

---

137. Victoria L Rubin, Niall J Conroy, Yimin Chen & Sarah Cornwell, "Fake News or Truth? Using Satirical Cues to Detect Potentially Misleading News" in *Proceedings of the Second Workshop on Computational Approaches to Deception Detection* (San Diego, 2016) 7.

138. Klonick, *supra* note 30 at 1651.

139. On the importance of introducing the level of intentions in the analysis of fake news see Naja Bentzen, "Understanding Propaganda and Disinformation" (2015), online (pdf): www.europarl.europa.eu [perma.cc/EPP7-F9T9]. See also *Report of the Special Rapporteur*, *supra* note 66; EC, *High-Level Expert Group on Fake News and Online Disinformation (HLEG): A Multi-Dimensional Approach to Disinformation* (2018), online (pdf): ec.europa.eu [perma.cc/4AYN-QLSL]. See also Organization for Security and Co-operation in Europe the Representative on Freedom of the Media (OSCE), *International Standards and Comparative National Approaches to Countering Disinformation in the Context of Freedom of the Media* (Vienna, March 2019).

140. See, generally, Donald A Barclay, *Fake News, Propaganda, and Plain Old Lies* (Rowman & Littlefield, 2018); see also A Brad Schwartz, *Broadcast Hysteria. Orson Wells's War of the Worlds and the Art of Fake News* (Farrar, Straus & Giroux, 2015).

141. Irini Katsirea, "'Fake News': Reconsidering the Value of Untruthful Expression in the Face of Regulatory Uncertainty" (2018) 10(2) J Media L 159 at 175.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

disinformation on a scale and with speed and precision of targeting that is unprecedented which create personalized information spheres and powerful echo chambers for disinformation campaigns";[142] b) the radical change of the European social context, which is now facing a massive crisis coming from "economic insecurity, rising extremism, and cultural shifts generate anxiety," and a general condition of "societal tensions, polarization, and distrust;"[143] c) the data that signals how traditional tools for fact-checking are not effective with web-expression;[144] and (d) the fundamental role that digital intermediaries (Facebook, Google, Twitter) have come to play in influencing the news media, by offering platforms where anyone can act as a publisher, and readers themselves—frequently "tricked" by clickbait titles—are invited to add comments and links to further contents.[145]

5. Taking Stock

This essay has explored the tests for balancing used by the ECtHR in matters of freedom of expression and proposed new categories for dealing with fake-news. It was suggested that the Court has usually relied on categorization and sliding-scales. The peculiarities of web-expression and the dangers connected to the enormous amount of bad quality information pose a new challenge. However, the legislative measures endorsed by the Member States and European institutions to counter "fake news" and disinformation on the Internet, while increasing "media literacy," might eventually yield illiberal consequences.

With an eye to the future, this essay argued that a minimal definition of "fake news", which incorporates the notion of intention, is fundamental for avoiding censorship in disguise and providing a conceptual framework for "constrained" balancing operations. Furthermore, the theoretical analysis has isolated a set of parameters that might be considered for weighing Art 10 in matters of web-expression. Finally, this inquiry has briefly addressed the opportunity of distinguishing fake news from traditional forms of false news since propaganda and misinformation are not new. However, new media, video-sharing, and bloggers are more powerful and effective means for the transmission of ideological contents.

---

142. *Tackling online disinformation*, *supra* note 67 at para 1; Syed, *supra* note 98 at 348ff. See also *Shtekel*, *supra* note 38 at para 63: "The risk of harm posed by content and communications on the Internet to the exercise and enjoyment of human rights and freedoms, particularly the right to respect for private life, is certainly higher than that posed by the press. Therefore, the policies governing reproduction of material from the printed media and the Internet may differ. The latter undeniably have to be adjusted according to the technology's specific features in order to secure the protection and promotion of the rights and freedoms concerned."

143. *Tackling online disinformation*, *supra* note 67 at para 2.2.

144. Council of Europe, Claire Wardle & Hossein Derakhshan, "Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking" (2017); Sam Levin, "Facebook Promised to Tackle Fake News. But the Evidence Shows it's not Working", *The Guardian* (16 May 2017), online: theguardian.com [perma.cc/A785-PK4Y].

145. Tarleton L Gillespie, "The Politics of 'Platforms'" (2010) 12(3) New Media & Society 347; Tarleton L Gillespie, "Governance of and by Platforms" in Jean Burgess, Alice Marwick & Thomas Poell, eds, *SAGE Handbook of Social Media* (Sage, 2018).

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press

The crusade against fake news should not become a hostage of the political preferences of national executives. Precise definitions are not “magic bullets”, but they can make a significant contribution to contrast illiberal forms of censorship to users, increase the rationality judicial balancing and the ability of online platforms to promote participation in EU democracies.

https://doi.org/10.1017/cjlj.2020.5 Published online by Cambridge University Press