

RESEARCH ARTICLE

Assessing plain and intelligible language in the Consumer Rights Act: a role for reading scores?

Kathy Conklin^{1†}, Richard Hyde^{2*} and Fabio Parente¹

¹School of English, University of Nottingham, Nottingham, UK and ²School of Law, University of Nottingham, Nottingham, UK

*Corresponding author. Email: richard.hyde@nottingham.ac.uk

(Accepted 27 March 2018)

Abstract

Under the Consumer Rights Act 2015 consumer contracts and consumer notices are required to be expressed in plain and intelligible language. This is a difficult concept to capture. Determining whether a contract is expressed in plain and intelligible language involves resource-intensive work by regulators and difficult adjudications by courts. This paper explores whether reading scores present a viable alternative. Can a simple computer program tell a consumer, a business, a regulator or the court that a particular contract is not expressed in plain and intelligible language? The paper begins by exploring the concept and role of plain and intelligible language in the Consumer Rights Act, before considering the ways that reading scores have developed and been used in legal contexts. We then report on the findings of an experimental examination of insurance contracts using a basket of reading scores, using our findings to draw conclusions about the utility of reading scores in determining whether a contract is expressed in plain and intelligible language. We find that reading scores can play a role in such determinations, but that further work is needed to provide appropriate tools for business, regulators and courts to use in assessing plain and intelligible language.

Keywords: Consumer Rights Act 2015; reading scores; transparency

Introduction

The requirement of transparency, which requires contracts to be drafted in plain and intelligible language and, if written, be legible, plays a key role in the governance of consumer contracts.¹ Businesses are incentivised to ensure plain and intelligible language in two ways. First, through protection from scrutiny of core terms expressed, inter alia, in plain and intelligible language.² Secondly, by the possibility of regulatory action (including an injunction) when a term or contract is not expressed in plain and intelligible language,³ whether or not the terms are substantively unfair. Therefore, it is necessary to have a clear understanding of the meaning of plain and intelligible language, and for business and regulators to operationalise the concept in order to assess whether contracts satisfy regulatory

[†]This work was supported by a Hermes Business Engagement Grant awarded by the University of Nottingham.

¹Transparency is also required by other legislative instruments. A key example is Regulation 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (GDPR), Art 12(1) (information should be provided in 'concise, transparent, intelligible and easily accessible form, using clear and plain language').

²HE Brandner and P Ulmer 'The Community Directive on Unfair Terms in Consumer Contracts: some critical remarks on the proposal submitted by the EC Commission' (1991) 28(3) Common Market Law Review 647 at 656.

³Consumer Rights Act 2015, s 68 and Sch 3, para 3(5), removing any doubts about the existence of such a power (see *Office of Fair Trading v Abbey National* [2008] EWHC 875 (Comm), [2008] 2 All ER (Comm) 625 at [86]).

requirements. Businesses can then draft compliant contracts, and regulators can take enforcement action against non-compliant ones.

Reading scores are an increasingly popular method for assessing contractual language and identifying contracts that are too difficult for the average consumer to understand. Advocates of the use of reading scores identify them as a 'simple, inexpensive way to measure the comprehensibility of legal language',⁴ something that is particularly useful when enforcement budgets are strained.⁵ Legislators, particularly in the United States, have used reading scores when seeking to reduce the complexity of legal documents, providing a clear target standard that documents must reach.⁶ Compliance with such standards can, in theory, be assessed with the application of a simple computer program.

This paper seeks to explore plain and intelligible language requirements in UK consumer contract law, and consider whether reading scores can (or cannot) provide a simple mechanism for determining whether a clause, or a document, is compliant with these requirements. It begins by considering in Part 1 the requirement that contracts be expressed in plain and intelligible language in its regulatory context, before turning to examine the development of reading scores in Part 2, examining some of the most common formulae used to determine readability. The use of reading scores to assess legal language is considered in Part 3, before Part 4 reports on an empirical test of the utility of reading scores in determining whether a contract is expressed in plain and intelligible language. Part 5 draws on the results of this empirical examination to make policy recommendations about the proper role of reading scores in the process of determining whether contract language is plain and intelligible.

1. Consumer contracts and plain and intelligible language

Standard form contracts dominate the relationship between consumers and traders. Consumers are not in a position to negotiate the terms of the contracts, and such contracts are often made on a 'take it or leave it' basis, giving businesses a high level of control of the terms of those contracts.⁷ The power imbalance has resulted in protection for consumers from the effects of the traders' ability to impose contract terms determined by them.⁸ This regulation of contract terms provides the background to the requirement that consumer contracts be expressed in plain and intelligible language.

(a) From contractual freedom to consumer contractual regulation

At common law there were few doctrines that allowed courts to interfere with the terms of a contract.⁹ Exclusion and limitation clauses, which eliminated or reduced the innocent party's entitlement to damages, were seen as particularly likely to give rise to unfairness.¹⁰ However, courts would not engage in the rewriting of contracts and would not refuse to enforce certain terms to prevent unfairness.¹¹ Interpretive rules were used to ameliorate harshness,¹² interpreting clauses in the manner most favourable to the party against whom they were used.¹³ These *contra proferentem* rules were perhaps the

⁴RW Benson 'The end of legalese: the game is over' (1984–85) 13 *Review of Law and Social Change* 519 at 547.

⁵See C Hodges *Law and Corporate Behaviour: Integrating Theories of Regulation, Enforcement, Compliance and Ethics* (Oxford: Hart, 2015) ch 14 part II.

⁶See below, text to n 107 et seq.

⁷See F Kessler 'Contracts of adhesion – some thoughts about freedom of contract' (1943) 43 *Columbia Law Review* 629 and AA Leff 'Contract as thing' (1970) 19 *American University Law Review* 131.

⁸See *Director of Fair Trading v First National Bank Ltd* [2001] UKHL 52, [2002] 1 AC 481 at [31] per Lord Steyn.

⁹The penalty jurisdiction is a notable exception. The development of the doctrine is summarised in *Cavendish Square Holding BV v Makdessi; ParkingEye Ltd v Beavis (ParkingEye)* [2015] UKSC 67, [2016] AC 1172.

¹⁰C Grunfeld 'Reform in the law of contract' (1961) 24 *MLR* 62 at 64.

¹¹See eg Denning LJ's statement in *Karsales (Harrow) Ltd v Wallis* [1956] 1 *WLR* 936 at 940 that the introduction of limitations on the scope of exclusion clauses was '[n]otwithstanding earlier cases which might suggest the contrary'.

¹²For example the 'rule' that prevented the application of exclusion clauses in cases of fundamental breach, disapproved by the House of Lords in *Suisse Atlantique Société d'Armement Maritime SA v NV Rotterdamsche Kolen Centrale* [1967] 1 AC 361.

¹³The *contra proferentem* rule(s) are retained by the Consumer Rights Act 2015, s 69(1).

closest common law precursor to the requirement to draft contracts in plain and intelligible language. Drafters were incentivised to ensure that the effect of their term was clear in order to ensure that the term would operate in the way they intended. Beyond interpretation, arguments were directed against the incorporation of such terms,¹⁴ with emphasis placed on the prominence of a term which could not be incorporated into the contract unless the party had notice of them.¹⁵

However, the common law was not thought to be sufficiently protective.¹⁶ In order to provide a legislative safeguard the Unfair Contract Terms Act 1977 (UCTA) was passed. This applied to both business-to-business and business-to-consumer contracts. Under UCTA courts could, in closely defined circumstances,¹⁷ adjudicate whether exclusion and limitation clauses¹⁸ were ‘reasonable’. UCTA did not separately regulate contractual wording, but the intelligibility of a term could be taken into account in assessing whether a term was reasonable.¹⁹

Further protection was provided for consumers by the Unfair Terms in Consumer Contracts Regulations 1994 (UTCCR 1994), which implemented the Unfair Terms Directive.²⁰ The UTCCR 1994 were repealed and replaced by the Unfair Terms in Consumer Contracts Regulations 1999 (UTCCR 1999). The primary function of the Unfair Terms Directive was to enable courts to decide whether terms were ‘fair’.²¹ This statutory intervention provided significantly improved protection for consumers, widening the scope of protection beyond exclusion and limitation clauses and allowing substantive issues to be canvassed. As a subsidiary matter, regulators and courts were required to assess whether core terms were expressed in plain and intelligible language when determining their jurisdiction, and given a power to assess all terms for linguistic compliance.

As part of the recent attempt to consolidate consumer contract law,²² rendering it less complex, less fragmented and clearer,²³ with respect to consumer contracts the Consumer Rights Act 2015 (CRA 2015) repealed and replaced UCTA and the UTCCR 1999.²⁴ UCTA continues to govern exclusion and limitation clauses in business to business contracts. Part 3 of the CRA 2015 implements the Unfair Terms Directive in a similar manner to UTCCR 1999. However, as the Directive requires only minimum harmonisation,²⁵ the UK has chosen to go beyond the minimum standards prescribed by providing that the CRA 2015 applies to individually negotiated contracts, imposing a requirement of prominence and legibility if a trader seeks to exempt a core term from scrutiny and applying the provisions to notices as well as contracts.

Under the CRA 2015 most of the terms in a consumer contract are subject to a test of fairness. The test is set out in CRA 2015, s 62(4). This provides that ‘[a] term is unfair if, contrary to the

¹⁴For example, cases such as *Parker v Southeastern Railway* (1877) 2 CPD 416; *Olley v Marlborough Court* [1949] 1 QB 532 and *Interfoto v Stilleto* [1989] QB 433.

¹⁵All terms in a signed contract will be incorporated (see *L'Estrange v Graucob* [1934] 2 KB 394).

¹⁶See Law Commission and Scottish Law Commission *Report 69: Exemption Clauses – Second Report* (Law Commission 1975) para [11].

¹⁷The term must fall within ss 2, 3 or 6, and the liability must relate to the course of a business or the occupation of the premises. A clause in a contract of the type set out in Sch 1 cannot be reviewed.

¹⁸The extended meaning given to such clauses covered by the Act are set out in s 13.

¹⁹See eg *Overseas Medical Supplier Ltd v Orient Transport Services Ltd* [1999] 2 Lloyd’s Rep 273 at 280.

²⁰Directive 93/13/EEC on unfair terms in consumer contracts.

²¹See Brandner and Ulmer, above n 2, and E McDonald ‘The emperor’s old clauses: unincorporated clauses, misleading terms and the Unfair Terms in Consumer Contracts Regulations’ (1999) 58(2) CLJ 413.

²²A consumer contract is ‘a contract between a trader and a consumer’ (CRA 2015, s 61(1)); ‘consumer’ and ‘trader’ are defined in CRA 2015, s 2(2).

²³CRA 2015, Explanatory Notes para [6].

²⁴For general consideration of the Act see D Barry et al *Blackstone’s Guide to the Consumer Rights Act* (Oxford: Oxford University Press, 2016); S Whittaker ‘Distinctive features of the new consumer contract law’ (2017) 133 LQR 47 and P Gilliker ‘The Consumer Rights Act 2015 – a bastion of European consumer rights?’ (2017) 37(1) Legal Studies 78.

²⁵See S Weatherill *EU Consumer Law and Policy* (Cheltenham: Edward Elgar, 2013) ch 3.

requirement of good faith, it causes a significant imbalance in the parties' rights and obligations under the contract to the detriment of the consumer'.²⁶ If a term is unfair it does not bind the consumer in his or her dealings with the trader.²⁷ The rest of the contract will continue, 'so far as practicable, to have effect in every other respect'.²⁸

The fairness test is not explored in detail in this paper. Instead, this paper focuses on the requirement that contractual terms be expressed in plain and intelligible language.²⁹ This requirement plays two roles in the scheme of the CRA 2015. First, (along with the requirements of prominence³⁰ and legibility³¹) it governs the subject matter scope of the s 62 fairness jurisdiction. If a term which relates to the main subject matter of the contract or the price payable is not expressed in plain and intelligible language then it may be subjected to scrutiny to assess whether it is fair. If such a term is expressed in plain and intelligible language, is legible and is prominent then it will be exempt from the test of fairness ('the core terms exception').³² The core terms exception seeks to encourage clarity in contract drafting.³³ Second, the requirement that a contract is expressed in plain and intelligible language is a standalone ground for regulatory action, which applies whether or not a contract term is fair.³⁴ This focuses on the drafting of the boilerplate terms of the contract rather than on the core terms.

(b) Plain and intelligible language

The requirement of plain and intelligible language is an important tool for ensuring that consumers are aware of the terms of trade.³⁵ Once a consumer is aware of the terms then he or she can make the choice not to engage with the trader on those terms,³⁶ forcing the trader to either change the term(s) or to exit the market.³⁷ Plain and intelligible terms are therefore an aid to market discipline,³⁸ and the requirement can be seen as a less interventionist approach to governance of unfair terms, as it empowers the consumer,³⁹ rather than interferes with the contract.⁴⁰ The linkage of the core terms 'empowerment' and 'plain and intelligible language' allows consumers to understand those terms that they must

²⁶The test for fairness was set out by the CJEU in *Aziz v Caixa d'Estalvis de Catalunya, Tarragona i Manresa* [2013] 3 CMLR 89, which was cited approvingly by the Supreme Court in *ParkingEye*, above n 9, where the key principles of the *Aziz* decision are set out at para [105]. The term in question was found to be fair.

²⁷CRA 2015, s 62(1).

²⁸CRA 2015, s 67.

²⁹CRA 2015, s 64(3).

³⁰A term is prominent if 'it is brought to the consumer's attention in such a way that an average consumer would be aware of the term' (CRA 2015, s 64(4)). An average consumer is 'reasonably well-informed, observant and circumspect' (see P Cartwright 'The consumer image within EU law' in C Twigg-Flesner (ed) *Research Handbook on EU Consumer and Contract Law* (Cheltenham: Edward Elgar, 2016)).

³¹Legibility is a UK law addition to the criteria contained in the Directive (Law Commission *Advice to the Department for Business, Innovation and Skills* (Law Commission, 2013) at para [4.3]), aimed at ensuring that the 'consumer [is] actually be given an opportunity to examine all the terms' (Unfair Terms Directive, recital 20). The legibility concept is beyond the scope of this paper.

³²CRA 2015, s 64.

³³Brandner and Ulmer, above n 2, at 656.

³⁴CRA 2015, s 68.

³⁵Information about the terms of trade is essential to the ability of consumers to make an informed choice in the market (see I Ramsay *Rationales for Intervention in the Consumer Marketplace* (OFT, 1984) para [3.8]).

³⁶Ramsay, *ibid*, at paras [3.26]–[3.28]; H Collins 'Good faith in European contract law' (1994) 14(2) *Oxford Journal of Legal Studies* 229 at 238, who states 'Clarity is essential for effective market competition between terms. What matters primarily for EC contract law is consumer choice...'

³⁷P Cartwright *Consumer Protection and the Criminal Law* (Cambridge: Cambridge University Press 2001) p 6.

³⁸In contrast see O Bar-Gill and O Ben-Shahar 'Regulatory techniques in consumer protection: a critique of European consumer contract law' (2013) 50 (Special Issue) CMLR 109 'people do not pay attention to standard forms...'

³⁹In contrast see S Weatherill 'Empowerment is not the only fruit' in D Leczykiewicz and S Weatherill (eds) *The Images of the Consumer in EU Law: Legislation, Free Movement and Competition Law* (Oxford: Hart, 2016) pp 212–214.

⁴⁰Brandner and Ulmer, above n 2, at 656.

read in order to make an informed choice,⁴¹ with the fairness test policing those terms that the average consumer does not read.⁴²

The Unfair Terms Directive requires ‘plain intelligible language’ in Art 4(2) and ‘plain, intelligible language’ in Art 5. This presents a slight contrast to CRA 2015, s 64(3), which requires ‘plain and intelligible’ language (authors’ emphasis). It is unclear whether the textual difference, which seems to make ‘plain’ and ‘intelligible’ separate criteria to be satisfied, will make any difference to the interpretation of the concept by the UK courts. This question arises particularly post-Brexit, where it appears that recourse to the CJEU to ensure consistency of interpretation may not be possible.⁴³ However, it does raise the possibility that separate tests and methodologies may be appropriate for determining whether a contract is plain and whether it is intelligible.

The concept of plain and intelligible language has been subject to a variety of interpretations. The CJEU has held that the requirement should be interpreted broadly, particularly because of the role it plays in excluding core terms from scrutiny.⁴⁴ Therefore, linguistic assessment ‘cannot be reduced merely to [a contract] being formally and grammatically intelligible’,⁴⁵ although a contract which fails to meet this test may not be seen as sufficiently ‘plain’. The assessment of language must also take into account whether the average consumer would be able to understand ‘potentially significant economic consequences’ of the term of the contract.⁴⁶ The requirement of plain and intelligible language therefore requires drafters to ensure that the effect of the term is communicated to the consumer. If the drafter fails to carry out this task the core term can be subject to scrutiny using the fairness test.⁴⁷

The mechanism for testing contracts for plain and intelligible language is underdeveloped. Courts have tended to make evaluative decisions based on the reading of the terms in their contractual context. In *Office of Fair Trading v Foxtons*,⁴⁸ it was held that ‘just because a highly skilled lawyer can find (or contrive) some equivocation in a word, that does not make the language lacking in plainness or intelligibility’⁴⁹ and that the assessment ‘does not require an absolute and pedantic rigour’.⁵⁰ However, a contract would not comply where it uses ‘broad terms of uncertain meaning... [w]ithout some form of definition’.⁵¹ This guidance leaves the determination to the judge, without fully fleshing out what is meant by plain and intelligible, particularly for the average consumer. In *Office of Fair Trading v Ashbourne Management Services* a term is seen as plain and intelligible if a consumer who reads the agreements ‘reasonably carefully’ can understand the effect of the terms.⁵²

In *Office of Fair Trading v Abbey National* Andrew Smith J decided that terms must be ‘sufficiently clear to enable the typical consumer to have a proper understanding of them for sensible and practical purposes’.⁵³ This does not explain how the assessment of this test is to be carried out. It is clear that non-contractual documentation available prior to the conclusion of the contract can be taken into account in determining whether a term was intelligible.⁵⁴ This has particular importance in complex

⁴¹This assumes that the consumer reads even the core terms of the contract. See Bar-Gill and Ben-Shahar, above n 38, and C Twigg-Flesner ‘The importance of law and harmonisation for the EU’s confident consumer’ in Leczykiewicz and Weatherill, above n 39, p 189.

⁴²M Chen-Wishart ‘Transparency and fairness in bank charges’ (2010) 126 LQR 157.

⁴³Department for Exiting the European Union ‘The United Kingdom’s exit from and new partnership with the European Union White Paper’ (DEEU, 2017) para [2.3].

⁴⁴See Case C-26/13 *Kásler and Káslerné Rábai* at [42].

⁴⁵Case C-96/14 *Van Hove v CNP Assurances SA* at [40].

⁴⁶*Ibid*, at [47].

⁴⁷Of course, a core term that is not transparent, but is fair, binds the individual consumer.

⁴⁸[2009] CTLC 188.

⁴⁹*Ibid*, at [62].

⁵⁰*Ibid*, at [63].

⁵¹*Ibid*, at [62].

⁵²*Office of Fair Trading v Ashbourne Management Services Ltd* [2011] CTLC 237 at [158].

⁵³*Office of Fair Trading v Abbey National*, above n 3, at [119].

⁵⁴*Ibid*, at [92].

financial transactions. Further, layout can be relevant to the assessment, with ‘useful headings and appropriate use of bold print’ making a contribution to the intelligibility of a document.⁵⁵

The role of the average consumer in the assessment of plain and intelligible language is under-explored.⁵⁶ Whilst the average consumer is specifically mentioned in the test of prominence,⁵⁷ it is not referred to in either the Unfair Terms Directive or the CRA 2015 when assessing plain and intelligible language.⁵⁸ At first instance in *Abbey National*, Andrew Smith J held ‘whether terms are in plain intelligible language is to be considered from the point of view of the ... average consumer. The... “average consumer ... who is reasonably well informed and reasonably observant and circumspect” ... provides an appropriate yardstick guide to whether a term is in plain intelligible language’.⁵⁹

However, the identity of the average consumer is not explored in the cases, with little attention given to the characteristics of such a consumer in the unfair terms context. In *Foxtons* there was little ‘evidence or other material to assist [the judge] in determining the mindset, thinking or attributes of a typical consumer’ and therefore assessment would be made ‘on an analogous footing to that on which the court approaches the attributes of the reasonable man in other realms’.⁶⁰ It is likely that the average consumer for whom language is assessed will be the average consumer in the targeted group. So if an insurance contract is targeted at a particular group of individuals, when assessing plain and intelligible language the average consumer will be the average consumer of that group.⁶¹ That the targeted standard applies in unfair terms cases is made clear in *Ashbourne Management Services*, where Kitchin J held that ‘[t]he question whether a particular term is expressed in plain intelligible language must be considered from the perspective of an average consumer. Here such a consumer is a member of the public interested in using a gym club which is not a high end facility and who may be attracted by the relatively low monthly subscriptions’.⁶²

Even if the average consumer is the benchmark against which plain and intelligible language is assessed, in contrast to the Consumer Protection from Unfair Trading Regulations 2008 (CPUTR 2008), no consideration is given to the proper consumer standard if a clearly identifiable group of consumers that is particularly vulnerable is foreseeably likely to enter into the contract.⁶³ The CPUTR 2008 provide that where a vulnerable group is ‘particularly vulnerable to the practice or the underlying product... in a way which the trader could reasonably be expected to foresee’ then ‘reference to the average consumer shall be read as referring to the average member of that group’.⁶⁴ The unfair terms case law does not seem to allow consumer vulnerabilities to be taken into account where the vulnerable group is not specifically targeted. The cases repeated reference to the ‘average consumer’⁶⁵ or the ‘typical consumer’⁶⁶ seem to suggest that a standardised approach will be adopted, without taking into account the reading skills of vulnerable consumers, even where they are foreseeably likely to enter into contracts on those terms.⁶⁷

⁵⁵Ibid, at [104].

⁵⁶Contrast the GDPR, above n 1, recital 58 (‘any information and communication, where processing is addressed to a child, should be in such a clear and plain language that the child can easily understand’) and Art 12(1). This is discussed in Growing Up Digital Taskforce *Growing Up Digital: A Report of the Growing Up Digital Taskforce* (Children’s Commissioner 2017) available at https://www.childrenscommissioner.gov.uk/wp-content/uploads/2017/06/Growing-Up-Digital-Taskforce-Report-January-2017_0.pdf (last accessed 4 November 2018) p 12.

⁵⁷See above n 30.

⁵⁸H Schulte-Nolke et al (eds) *EC Consumer Law Compendium: A Comparative Analysis* (2008) p 398.

⁵⁹*Office of Fair Trading v Abbey National plc*, above n 3, at [89].

⁶⁰Above n 48, at [31].

⁶¹Made explicit in the Consumer Protection from Unfair Trading Regulations 2008 (CPUTR), reg 2(4).

⁶²Above n 52, at [155].

⁶³In the CPUTR 2008, vulnerability exists on account of ‘mental or physical infirmity, age or credulity’ (reg 2(5)).

⁶⁴Ibid.

⁶⁵*Ashbourne Management Services*, above n 52, at [155].

⁶⁶*Office of Fair Trading v Abbey National plc*, above n 3, at [119]; *Foxtons*, above n 48, at [29].

⁶⁷It may be possible that terms that are particularly damaging to vulnerable consumers are misleading actions under CPUTR 2008.

In the light of the failure of the UK courts to flesh out the concept, the Competition and Markets Authority has offered an account of the characteristics of a plain and intelligible document.⁶⁸ This guidance is, amongst other things, intended to assist business in drafting contracts, protecting their core terms from scrutiny and ensuring consumers can make an informed choice.⁶⁹ First, the document should be jargon free, and should ‘as far as possible use ordinary words in their normal sense’. This will not prevent the use of technical language where the meaning of the language is clear to the consumer.⁷⁰ However, such statements raise the question of who the consumer is, what is clear to them, and whether some technical language can ever be clear. Secondly, the document should be unambiguous, meaning ‘clear and not open to misinterpretation or differing interpretations’; thirdly, it should be reader-friendly, including ‘organised so as to be easily understood (using, for example, short sentences and subheadings)’; fourthly, it must be comprehensible with ‘the meaning of the words or concepts used, as well as the reasons for them... explained if they are not capable of being readily understood by consumers’; fifthly, it should be informative, so that ‘a consumer should, on the basis of the information provided – if necessary in pre-contractual literature – be able to foresee and evaluate the consequences of all wording used’;⁷¹ and sixthly, accompanied by pre-contractual literature as necessary, for example ‘if, for instance, the contract is complex or lengthy’.

Whilst providing a useful typology, and whilst such a multifactorial approach is useful allowing decision-makers to take into account various different matters in judging whether language is plain and/or intelligible, the concepts used in the typology suffer from a lack of theoretical or empirical grounding. They do not, by themselves, provide a simple way of measuring whether a clause or a contract is transparent, but instead provide a series of matters that a decision-maker may take into account in making an evaluative judgment about whether a contract is expressed in plain and intelligible language. It is accepted that providing a simple measure is a difficult task, but it is one that should be undertaken in an attempt to ensure that the important concept of plain and intelligible language is sufficiently certain. If not, there may be uncertainty for businesses, who cannot easily judge whether their contracts are compliant, and difficulty for consumers, who are unable to easily decide whether the core terms are challengeable or not, particularly in circumstances where they are not legally advised. Further, it is more expensive to conduct a multifactorial approach than to have a simple metric that can determine whether a contract is expressed in plain and intelligible language. Reading scores, examined in Parts 2 to 4, are an attempt to simplify the assessment of contractual language and reduce the time and expense of a multifactorial approach.

The problems caused by this broad multifactorial approach are exacerbated by the general requirement of plain and intelligible language imposed in the CRA 2015. When assessing the applicability of the core terms exception, the linguistic examination is limited to terms governing price and subject matter. However, s 68 provides that a trader must ensure that a ‘written term of a consumer contract... is transparent’. Therefore, all terms in consumer contracts must be expressed in plain and intelligible language. It is questionable whether the concept of plain and intelligible language is subject to the same test under s 68 as that set out in core terms exemption. Boilerplate terms may not have ‘significant economic consequences’,⁷² but may significantly affect the rights and obligations of the parties to the contract.⁷³ It is, therefore, necessary to examine whether the term as written allows the contracting parties to understand this.

Where a contract is not transparent, regulators may take action under CRA 2015, Sch 3 and apply for an injunction (or accept an undertaking) that prevents a trader using that term. The injunction jurisdiction is necessary as ‘one cannot think of a more expensive or frustrating course than to seek

⁶⁸Competition and Markets Authority *Unfair Contract Terms Explained* (CMA, 2015) para [40].

⁶⁹Above, text to n 35 et seq.

⁷⁰For example in consumer insurance contracts (see J O’Sullivan and J Hilliard *The Law of Contract* (Oxford: Oxford University Press, 7th edn, 2016) p 217).

⁷¹Reflecting the decision in *Van Hove*, above n 45.

⁷²Of course, consequences of, for example, exclusion and limitation clauses may be significant in economic terms.

⁷³Eg an arbitration clause.

to regulate... “contract” quality through repeated lawsuits against inventive “wrongdoers”,⁷⁴ and consumers either cannot or will not take private law actions.⁷⁵ By giving power to regulators to take enforcement action to prevent the use of unclear contracts, regulation seeks to provide a better mechanism to improve the quality of contracts compared to consumer civil actions.⁷⁶ However, in times of austerity,⁷⁷ the enforcement of the linguistic requirements is likely to be limited. Regulators will only take action in high impact cases.⁷⁸ A compliance-focused, responsive regulation,⁷⁹ approach is likely to be adopted, with engagement with the trader in order to co-operatively lead to changed, plain and intelligible, terms. If a simple test could be developed, this would allow regulators to engage with lack of compliance on a more regular basis, providing a simple starting point to negotiated changes to the terms used by businesses.⁸⁰

In an attempt to develop a better understanding of the concept of plain and intelligible language, we considered whether reading scores could provide an accurate test of whether a document was expressed in plain and intelligible language. If this could be determined by a reading score it would be extremely helpful to consumers, regulators and traders, as documents could be scrutinised in an efficient and cost-effective manner using a simple computer program.⁸¹ We chose to examine the exclusions in consumer insurance contracts. Whilst these exclusions would initially appear to be subject a fairness test and a s 68 assessment, in consumer insurance contracts these clauses are core terms as they define the subject matter (the scope of the insurers’ risk) and the price (as they contribute to ‘calculating the premium paid by the consumer’).⁸² Therefore, in the rest of this paper we are considering the contribution that reading scores can make to the assessment of core terms, although, where necessary, we discuss the potential of reading score in assessing whether boilerplate is expressed in plain and intelligible language. Our findings are set out in the following sections.

2. The development of reading scores

Reading scores attempt to quantify the ease with which readers are able to read and comprehend written texts.⁸³ Such scores have long been used to assess a variety of texts. They seek to provide a simple measure of how readable a piece of text is, in order that a writer can assess and amend the text to make it understandable to readers.⁸⁴

In one of the earliest attempts, Thorndike compiled a list of 10,000 words occurring in general literature by frequency of use, suggesting that the readability of written texts could be determined mathematically.⁸⁵ Thorndike’s list served as partial basis for one of the earliest readability formulae,

⁷⁴AA Leff ‘Unconsonability and the crowd: consumers and the common law tradition’ (1970) 31 *University of Pittsburgh Law Review* 349 at 356.

⁷⁵H Collins *Regulating Contracts* (Oxford: Oxford University Press, 1999) pp 87–93.

⁷⁶See I Ramsay *Consumer Law and Policy: Texts and Materials on Regulating Consumer Markets* (Oxford: Hart, 3rd edn, 2012) pp 317–320.

⁷⁷Hodges, above n 5, ch 14 part II.

⁷⁸Such as *Office of Fair Trading v Abbey National*, above n 3, and *Foxtons*, above n 48.

⁷⁹I Ayres and J Braithwaite *Responsive Regulation: Transcending the Deregulation Debate* (Oxford: Oxford University Press, 1992); and Cartwright, above n 37, pp 220–222.

⁸⁰Ramsay, above n 76, pp 320–330 and S Bright ‘Winning the battle against unfair contract terms’ (2000) 20(3) *Legal Studies* 331.

⁸¹The ease of reading scores as a tool for assessing compliance is noted by Benson, above n 4.

⁸²Directive 93/13/EEC, recital 19.

⁸³JS Chall and E Dale *Readability Revisited: The New Dale-Chall Readability Formula* (Brookline Books, 1995) pp 79–80 define a readability formula as ‘an equation which combines those text features that best predict text difficulty. The equation is usually developed by studying the relationship between text features (eg words, sentences) and text difficulty (eg reading comprehension, reading rate and expert judgment of difficulty)’.

⁸⁴Such is the function of the reading scores built in most word processing programs (see for example <https://support.office.com/en-gb/article/Test-your-document-s-readability-85b4969e-e80a-4777-8dd3-f7fc3c8b3fd2> (last accessed 4 November 2018).

⁸⁵EL Thorndike *The Teacher’s Word Book* (Teacher’s College, Columbia University, 1921).

published by Vogel and Washburne, known as the Winnetka formula.⁸⁶ It considered factors such as the number of different words per 1000 words, the number of uncommon words (words not on Thorndike's list) per 1000 words, the number of prepositions per 1000 words, and the number of simple sentences in 75 successive sentences. Readability scores were computed for passages from 700 books and validated against children's paragraph-meaning scores (a measure of reading comprehension of those passages).

Later work conducted by Waples and Tyler⁸⁷ and Ojemann⁸⁸ explored other factors beyond word frequency that may influence readability, and the following three decades saw the development of a number of readability measures whose subsequent revisions are still in widespread use today. Flesch developed the Flesch reading-ease score (FRES),⁸⁹ which assigns texts a numerical score between 0 and 100, with lower scores indicating more difficult texts. The formula uses the ratio between the total number of words and the total number of sentences, and the ratio between the total number of syllables and the total number of words. It is calculated as follows:

$$FRES = 206.835 - 1.015 * \left(\frac{\text{total words}}{\text{total sentences}} \right) - 84.6 * \left(\frac{\text{total syllables}}{\text{total words}} \right)$$

The formula was subsequently revised with different weights to produce a score interpretable as the target US grade level of the text, or the number of years of formal schooling required to understand its content.⁹⁰ The Flesch-Kincaid (F-K) grade level is computed as follows:

$$F-K = 0.39 * \left(\frac{\text{total words}}{\text{total sentences}} \right) - 11.8 * \left(\frac{\text{total syllables}}{\text{total words}} \right) - 15.59$$

Another popular formula, the Gunning FOG Index, was created by Robert Gunning,⁹¹ and is computed as follows:

$$FOG = 0.4 * \left[\left(\frac{\text{words}}{\text{sentences}} \right) + 100 * \left(\frac{\text{complex words}}{\text{words}} \right) \right]$$

where 'complex words' are defined as words containing three or more syllables. McLaughlin subsequently formulated an alternative to the FOG Index,⁹² called the SMOG Index, which produced a grade level estimate using the following formula:

$$SMOG = 1.0430 * \sqrt{\text{number of polysyllables}} * \frac{30}{\text{number of sentences}} + 3.1291$$

The measures described thus far include the number of syllables as a factor. However, the syllable structure of English is quite complex and varied, making automation of syllable counting problematic. To address this, two measures were developed to expedite the automation of readability computations, and were based on the number of characters (letters) per word instead of syllables. These are the

⁸⁶M Vogel and C Washburne 'An objective method of determining grade placement of children's reading material' (1928) 28(5) *The Elementary School Journal* 373.

⁸⁷D Waples and RW Tyler *What People Want to Read About* (University of Chicago Press, 1931).

⁸⁸R Ojemann 'The reading ability of parents and factors associated with the reading difficulty of parent education materials' (1934) 8 *University of Iowa Studies in Child Welfare* 11.

⁸⁹R Flesch 'A new readability yardstick' (1948) 32(3) *Journal of Applied Psychology* 221.

⁹⁰JP Kincaid et al *Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel (No RBR-8-75)* (Naval Technical Training Command, 1975).

⁹¹R Gunning *The Technique of Clear Writing* (McGraw-Hill, 1952).

⁹²GH McLaughlin 'SMOG grading - a new readability formula' (1969) 12(8) *Journal of Reading* 639.

Automated Readability Index (ARI)⁹³ and the Coleman-Liau Index (CLI)⁹⁴:

$$ARI = 4.71 \left(\frac{\text{characters}}{\text{words}} \right) + 0.5 \left(\frac{\text{words}}{\text{sentences}} \right) - 21.43$$

$$CLI = 0.0588 * L - 0.296 * S - 15.8$$

where L is the average number of characters per 100 words, and S the average number of sentences per 100 words.

Considerable debate exists as to the relative merits of formulae based on characters versus syllables, but ARI and CLI owe much of their popularity to the relative ease and reliability with which computer programs can compute the number of characters of English words as opposed to the number of syllables they contain. Syllable-based measures computed by different analytical tools can produce more or less significant discrepancies depending on the particular syllable parsers (computer programs that identify the syllables in the words assessed, in order that the number of syllables can take their place in the formula) employed and the specific assumptions these make. For example, some syllable parsers might treat ‘fine-tuning’ as two words, and others as one. These differences during parsing can have a noticeable impact on the grade level estimate. Marchand, Adsett and Damper provide a useful overview and comparison of different rule-based and data-driven syllabification algorithms.⁹⁵ It is important to note that there is no uniform methodology for parsing, and therefore a clear risk that different parsers come to different outcomes.

Beyond those examined in this section, hundreds of different readability measures have been formulated (for English and for other languages), and a review of all of them is beyond the purview of this paper. Despite the large number of readability measures, all the formulae described in this section remain in widespread use. One of the uses that the reading scores have been put to is examining legal documents. Before using these formula to examine sample contracts, it is useful to consider the academic and legislative uses of reading scores to scrutinise legal language, and particularly contracts.

3. The use of reading scores in assessing legal language

Reading scores are increasingly being used in legal scholarship and legal policy, often as an aid to challenges to ‘legalese’, which is seen as too arcane and difficult to understand.⁹⁶ The linguistic requirement in the CRA 2015 seeks to challenge such language. Therefore, it is natural to consider whether reading scores can help the requirement of plain and intelligible language achieve its policy goal.

Academically, a number of studies have examined legal, or quasi-legal, documents for readability. Reading scores have been used to examine tax legislation in Australia⁹⁷ and New Zealand.⁹⁸ Sutherland used the Flesch Reading Ease and Flesch-Kincaid Grade Level formulae to examine collective bargained agreements in Australia,⁹⁹ using the results to assess whether the policy goal of ensuring

⁹³EA Smith and RJ Senter ‘Automated readability index’ (1967) AMRL-TR. Aerospace Medical Research Laboratories (6570th) 1.

⁹⁴M Coleman and TL Liu ‘A computer readability formula designed for machine scoring’ (1975) 60(2) *Journal of Applied Psychology* 283.

⁹⁵Y Marchand et al ‘Automatic syllabification in English: a comparison of different algorithms’ (2009) 52(1) *Language and Speech* 1.

⁹⁶Benson, above n 4.

⁹⁷D Smith and G Richardson ‘The readability of Australia’s taxation laws and supplementary materials: an empirical investigation’ (1999) 20 *Fiscal Studies* 321.

⁹⁸A Sawyer ‘New Zealand’s tax rewrite program – in pursuit of the (elusive) goal of simplicity’ (2007) 4 *British Tax Review* 405, and the articles cited therein.

⁹⁹C Sutherland ‘The elusive quest for simplicity: measuring and assessing the readability of enterprise agreements, 1993 to 2011’ (2013) 35 *Sydney Law Review* 349.

'simple agreements' had been achieved. Rogers et al have used reading scores to analyse the readability of Miranda¹⁰⁰ warnings given to suspects in the USA.¹⁰¹

In the UK, reading scores have been used to examine consumer contracts, including those on the internet.¹⁰² Linsley and Lawrence examined the risk disclosure section of annual reports of Public Listed Companies using Flesch Reading Ease scores, and found the disclosures were difficult or very difficult to understand.¹⁰³ On the policy front, in their examination of children's digital lives, the Growing Up Digital Taskforce used a Flesch-Kincaid reading score to evaluate the terms and conditions of Instagram.¹⁰⁴ The terms were found to be 'difficult to read' with 'language and sentence structure that only a postgraduate could be expected to understand'.¹⁰⁵ Consumer groups have used reading scores in an attempt to encourage traders to redraft their contracts. For example, Fairer Finance found that 'the average insurance document is only accessible to someone in the last year of Sixth Form College'.¹⁰⁶

The use of reading scores is embedded in some US legislation,¹⁰⁷ requiring contracts (including insurance contracts) to have a particular level of readability.¹⁰⁸ For example, in Texas a consumer banking contract which is not concluded on model contract provisions must meet prescribed Flesch-Kincaid reading scores, as calculated by Microsoft Word.¹⁰⁹ Similarly, the South Carolina code requires that loan contracts have a Flesch-Kincaid score of 'no higher than seventh grade'.¹¹⁰ The Montana code provides that an insurance policy cannot be issued in Montana unless 'the text achieves a minimum score of 40 on the Flesch reading ease test'.¹¹¹ There have been some moves towards the use of reading scores in financial documentation in Canada, but legislative action has been stalled because French language reading scores are not felt to be sufficiently developed to provide an appropriate legislative benchmark.¹¹² Sirico is critical of these developments, arguing that the legislation requiring the use of reading scores does not protect consumers.¹¹³ He further argues that the calculation of Flesch-Kincaid scores on Microsoft Word is not in accordance with the standard formula, as it fails to count syllables, and instead uses number of characters as a substitute.¹¹⁴ This has particular implications for the Texas statute discussed above.

¹⁰⁰*Miranda v Arizona* (1966) 384 US 436.

¹⁰¹R Rogers et al 'The language of Miranda warnings in American jurisdictions: a replication and vocabulary analysis' (2008) 32 *Law and Human Behavior* 124.

¹⁰²M Hochhauser 'Compliance vs communication' (2003) 50 *Clarity: Journal of International Movement to Simplify Language* 11; S Moran, E Luger and T Rodden 'Literatin: beyond awareness of readability in terms and conditions' (2014) *Proceedings of Ubicomp* '14.

¹⁰³PM Linsley and MJ Lawrence 'Risk reporting by the largest UK companies: readability and lack of obfuscation' (2007) 20 *Accounting, Auditing & Accountability Journal* 620.

¹⁰⁴*Growing Up Digital*, above n 56.

¹⁰⁵*Ibid*, p 8.

¹⁰⁶See Fairer Finance 'Insurance and banking customers need a PhD to understand the small print' available at <https://www.fairerfinance.com/insights/press-releases/insurance-and-banking-customers-need-a-phd-to-understand-the-small-print>.

¹⁰⁷See in general C Felsenfeld 'The plain English movement' (1982) 6 *Canadian Business Law Journal* 408.

¹⁰⁸See JA Cogan Jr 'Readability, contracts of recurring use, and the problem of ex post judicial governance of health insurance policies' (2010) 15(1) *Roger Williams University Law Review* 93 at 120 fn 105 and see generally MS Friman 'Plain English statutes' (1995) 7 *Loyola Consumer Law Reporter* 103; R Pressman *Legislative and Regulatory Progress on the Readability of Insurance Policies* (Document Design Center, 1979) and CJ Karlin 'Readability statutes – a survey and a proposed model' (1980) 28 *University of Kansas Law Review* 531.

¹⁰⁹Texas Administrative Code title 7 section 90.1045.

¹¹⁰Code of Laws of South Carolina 1976 title 37 chapter 3 section 202.

¹¹¹Montana State Code title 33 chapter 15 section 325. The Montana code provides detailed methodology for use in calculating the Flesch reading ease score.

¹¹²See eg Ontario Securities Commission Bulletin Issue 33/40s4 (October 08, 2010) Appendix A available at http://www.osc.gov.on.ca/en/SecuritiesLaw_rule_20101008_81-101_pos-oscb_3340-sup-4.htm (last accessed 4 November 2018).

¹¹³LJ Sirico Jr 'Readability studies: how technocentrism can compromise research and legal determinations' (2008) 26(1) *Quinnipiac Law Review* 147.

¹¹⁴*Ibid*, at 165–166.

Given the trend towards using reading scores to assess contracts, it is timely to consider whether such scores could, and should, be used in the assessment of plain and intelligible language. In the next part we apply the reading scores considered in Part 2 to consumer insurance contracts in order to consider their utility in assessing the linguistic compliance of contracts.

4. Testing insurance contracts

Textual extracts of general exclusion clauses from seven different consumer travel insurance documents were scrutinised.¹¹⁵ These seven insurance policies were selected randomly from a population of consumer travel insurance contracts selected by the authors using a price comparison website. The authors made four separate searches and harvested the policy wordings returned in response.¹¹⁶ The four searches were as follows: a two-week break in Europe;¹¹⁷ a two-week break outside Europe (excluding US and Canada);¹¹⁸ a two-week break in the US;¹¹⁹ and an annual travel insurance policy.¹²⁰ A number of policy wordings were returned multiple times in response to each of the searches, and where the wording of the exclusions was the same these were removed from the population. Each exclusion wording was assigned a number and seven were randomly selected. The policies analysed were all issued by different companies. The analysed sections ranged between 568 and 1747 words in length. Where necessary, texts were formatted to remove numbered/alphabetised lists and bullet points, but leaving the sentences otherwise intact (including numbers when embedded in sentences).

The goal of this examination was to achieve a better understanding of the utility of reading scores for assessing the concept of plain and intelligible language. We also attempted to test if the reading scores measured consumer understanding using questioning about the effect of the terms contained in the contract. Consumers were provided with the contracts and were given a series of 28 vignettes relating to losses incurred during a holiday,¹²¹ and asked whether the insurance responded to the risk. Two example vignettes are set out in Figure 1. The answers given were then coded as correct or incorrect, with incorrect answers indicating consumer inability to understand the effects of the term. The findings are used to draw conclusions about the usefulness of reading scores in assessing whether language is plain and intelligible, and whether they should be used by courts, regulators and traders.

Consumer insurance contracts were selected as the subjects of the study for a number of reasons. First, plain and intelligible language, and the protection from scrutiny under the CRA 2015 it affords, is particularly important in the insurance industry as both insuring clauses and exclusions are 'core terms'.¹²² Collins suggests that this broad conception of core terms 'threatens to exempt insurance contracts from control by the back door'.¹²³ This means that the conceptualisation of plain and intelligible language is particularly important for consumers and insurers, as large parts of insurance contracts can be protected from substantive scrutiny by virtue of the core terms exception.¹²⁴ Therefore, a well-developed concept of plain and intelligible language is particularly necessary in this area.

Second, consumer insurance contracts are pervasive. Most consumers will hold insurance against some risks.¹²⁵ The wording of insurance contracts can have important implications for consumers' entitlements. The harshness of the contractual position, that consumers are not entitled to an

¹¹⁵Documents on file with the authors.

¹¹⁶For the purpose of the search the traveller was assumed to have no pre-existing conditions, and did not require any special insurance covering winter sports or business travel.

¹¹⁷n = 113.

¹¹⁸n = 108.

¹¹⁹n = 111.

¹²⁰n = 112.

¹²¹A Bryman *Social Research Methods* (Oxford: Oxford University Press, 2nd edn, 2004).

¹²²Directive 93/13/EEC, recital 19 provides 'the terms which clearly define or circumscribe the insured risk and the insurer's liability shall not be subject to [a fairness] assessment'.

¹²³H Collins 'Good faith in European contract law' (1994) 14(2) *Oxford Journal of Legal Studies* 229 at 243.

¹²⁴See eg *Bankers Insurance Co Ltd v South* [2003] EWHC 380 (QB), [2003] PIQR P28.

¹²⁵For example, buildings and contents insurance, car insurance, travel insurance, pet insurance, gadget insurance, etc.

Example vignette 1: Ben is travelling with his friend Steve. Steve has been taking heavy antibiotics and had been advised to cancel his trip. A few days into the trip, his condition worsens to the point of requiring emergency medical care and causing Ben to have to cancel the rest of the holiday. Will Ben be able to claim compensation under this policy?

Example vignette 2: Sarah has booked a holiday in Kenya, but, shortly before she is due to fly there, the holiday resort is shut down due to civil unrest in the region and credible bomb threats. Will Sarah be able to claim her money back under this policy?

Figure 1. Example vignettes

indemnity in the event that they fall within exclusions, has been to some extent ameliorated by the availability of the Financial Ombudsman Service,¹²⁶ who have shown a willingness to uphold some complaints about the application of an exclusions where a matter appears to fall strictly within the policy wording.¹²⁷ However, it is still the case that consumers can often find themselves without compensation after suffering injury due to the terms of an exclusion in an insurance contract. Further, travel insurance is likely to be the most complex financial product that consumers buy on a year-to-year basis, and therefore provides a useful case study for considering reading scores.

We subjected each extract to analysis using a number of different reading score formulae. We computed the readability measures described above (F-K, FOG, SMOG, ARI, and CLI) for our seven extracts. The readability measures were computed for each text using six different analytical tools/calculators. These included the koRpus R package for text analytics (version 0.06-5);¹²⁸ using the English version of the TreeTagger parser (version 3.2.1);¹²⁹ an online implementation of koRpus; and four other freely available online tools capable of computing the required readability measures.

There are differences in the reading scores returned by the different indicators. This can be seen in Table 1, which shows the different scores returned by the different indicators using each different method of calculating the scores for one of the insurance documents examined during the project. The reading scores calculated vary from around 13 years of education using CLI to almost 20 of education using FOG. As seen above in Part 2, each indicator is examining different characteristics of the contract, and therefore each provides potentially valuable information on readability. By choosing one indicator, the complexity measured by other indicators is lost, and therefore documents that may be challenging in a way captured by a particular score will not be identified if a different reading score is chosen as the measure of plain and intelligible language.¹³⁰

Further, it was noticed that there were differences between the scores calculated by different calculators. For example, the calculations of FOG have a range for years of education that differs by 4.59. This is due to potential (and not necessarily transparent) differences between parsing algorithms for syllables used by different analytical calculators. This is potentially problematic for regulators, businesses and consumers, as although a reading score may be perceived to be fixed, it clearly is not. This means that the legality of contract terms, or a decision to take enforcement action, may be dependent on which calculator is used. This is particularly problematic if a stakeholder uses a reading score as a proxy for plain and intelligible language, where the same document can appear to be

¹²⁶See generally W Merricks 'The Financial Ombudsman Service: not just an alternative to court' (2007) 15(2) *Journal of Financial Regulation and Compliance* 135.

¹²⁷See eg Ombudsman News Issue 29 Case 29/1.

¹²⁸M Eik Michalke *koRpus: An R Package for Text Analysis (Version 0.06-5)* (2016), available at <http://reaktanx.de/?c=hacking&s=koRpus> (last accessed 4 November 2018).

¹²⁹H Schmid 'Improvements in part-of-speech tagging with an application to German' (1995) *Proceedings of the ACL SIGDAT-Workshop*.

¹³⁰This is a particular risk in the legislative approach taken in those States in the USA where one formula is used (see above text to n 107 et seq).

Table 1. Reading score in years of education by formulae and calculator for travel insurance contract 1

		Reading Score Formulae					<i>Mean</i>	<i>Standard Deviation</i>
		Automated Readability Index	Coleman-Liau Index	Flesch-Kincaid	Gunning FOG Index	SMOG		
Reading Score Calculators	koRpus (R)	14.81	12.73	14.47	16.49	16.82	15.06	1.65
	koRpus (online)	16.27	13.22	16.88	21.08	18.21	17.13	2.86
	Readability Consensus Calculator	16.9	12	17	20.3	15.4	16.32	3.0
	OnlineUtility	17.11	12.99	17.62	19.7	18.13	17.11	2.49
	Perry Marshall	16.9	13.8	17	20.4	15.5	16.72	2.43
	Readability Test Tool	16.9	13.8	17	20.4	15.5	16.72	2.43
<i>Mean</i>		16.48	13.09	16.66	19.72	16.59		
<i>Standard Deviation</i>		0.86	0.68	1.10	1.64	1.32		

transparent when using one calculator, but not if another is used. One solution is to be prescriptive as to the methodology by which the score will be calculated, favouring one calculator and/or method of calculation.¹³¹ However, such a prescription has the potential to embed problematic calculation errors within the concept of plain and intelligible language, with the score produced by a chosen calculator leading to a conclusion that a document does or does not reach a set reading score threshold that is used to determine that it is plain and/or intelligible, and therefore is or is not transparent.

Therefore, it is argued that a better approach would be to pool grade-level estimates produced by different measures and calculators together to produce a consensus as to the readability of each text under analysis. Pooling these estimates would produce a single Grand Weighted Mean grade level for each text under investigation. We suggest that such an approach has the potential to eliminate some of the measurement issues that arise in the legislative and regulatory use of reading scores. Such an approach has not been used by stakeholders, who often rely on one formula or one calculator or both. In order to produce the Grand Weighted Mean, each calculator was used to produce all five of the readability measures (ARI, CLI, FK, FOG, SMOG) for each text. Then, for each text, grand means were computed for each measure across the different calculators, and for each calculator across different measures. Greater weight was assigned to those measures and calculators providing more consistent results and less weight to those yielding greater variance. The process was repeated to obtain a single Grand Weighted Mean grade-level estimate of readability for each text. This is a novel approach, taking publically available tools and utilising them, whilst attempting to resolve inconsistencies in calculation. This approach has high utility for users such as regulators, business and consumers, who are not equipped to engage in ad hoc calculation of reading scores or to create bespoke 'better' reading score methodologies,¹³² but who wish to utilise widely available tools to assess contractual language.

The Grand Weighted Means are set out in Table 2. Testing indicates that more than 19 years of formal education (beyond a Master's Degree) is required to understand Legal Text 2, while Legal Text 5 needs just under 14 years (second year of university). Despite Text 5 achieving the lowest reading score, none of our texts appear to be particularly plain and intelligible. Unless the average consumer (if involved in the linguistic analysis) were conceptualised as having the reading ability of at least a second year university student then it would appear that, in a system that assessed compliance by utilising reading scores alone, then all of our insurance texts would be seen as lacking transparency, and therefore subject to scrutiny under the fairness test. Rather than Collins' concern of 'exempt[ing] insurance contracts from control',¹³³ such a test of plain and intelligible language has the potential to bring exclusions within the fairness regime, seemingly contrary to the intention of the drafters of the Unfair Terms Directive.¹³⁴ It may be that analysis at clause level would produce some clauses that achieve acceptable reading scores, whilst others do not, but such a close focus is likely to reduce any efficiency gains of using reading scores. Therefore, it appears that reading scores may not provide an appropriate method for assessing plain and intelligible language, at least for core terms in insurance contracts, as they may bring terms within the fairness test which were intended to be excluded. However, it may be that exclusions drafted like those in our sample should be subject to the test for fairness, in order to encourage more transparent drafting and protect the consumer.

Following the application of the reading scores it was necessary to consider whether reading scores reflect the intelligibility of the contracts. One way of doing this is to consider whether consumers understand the terms of contracts better when the reading score is more favourable. A study that looks at intelligibility is Davis.¹³⁵ In one experiment he examined the ability of consumers to

¹³¹As in the Montana State Code, above n 111.

¹³²Creating a new reading score methodology tends to be the response in the linguistics literature when identifying consistency problems with the scores or the calculators (see, for example, M Sinha 'New readability measures for Bangla and Hindi texts' (2012) Proceedings of COLING 2012: Posters 1141, 1144).

¹³³Collins, above n 123.

¹³⁴Directive 93/13/EEC, recital 19.

¹³⁵J Davis 'Protecting consumers from overdisclosure and gobbledygook: an empirical look at the simplification of consumer-credit contracts' (1977) 63(6) Virginia Law Review 841.

Table 2. Mean years of education required to comprehend each contract

Text	Grand Weighted Mean (years of education)
Travel Insurance Contract 1	16.29
Travel Insurance Contract 2	19.06
Travel Insurance Contract 3	15.95
Travel Insurance Contract 4	16.35
Travel Insurance Contract 5	13.86
Travel Insurance Contract 6	14.05
Travel Insurance Contract 7	15.06

comprehend a contract. The contract was either an unchanged version of a consumer credit contract or was a redrafted version, simplified (reducing unnecessary clauses to reduce the possibility of information overload¹³⁶) and amended for readability. The second group, who read the redrafted contract, scored 26% better on the test of their understanding. Vulnerable consumers (young and/or poor and/or African-American and/or inexperienced) showed the greatest improvement in their score when using the redrafted contract. Further experimentation showed that both simplification and redrafting for readability were necessary to achieve the increased understanding.

Further studies on comprehension have found that contractual simplification has a positive effect for consumers. Masson and Waldron¹³⁷ redrafted a contract by removing redundant or archaic terms, simplifying words and sentence structures and defining or simplifying legal terms. Comprehension, as measured by paraphrasing and question-answering tasks, was reliably and significantly enhanced by the use of simplified words and sentence structure. However, absolute levels of comprehension were still low. Further, defining and or simplifying legal terms did not have a significant impact on comprehension.¹³⁸

Using the vignettes to assess comprehension of the consequences of the terms, we found that lower reading scores did not necessarily translate to improved understanding. There was no significant correlation between reading score and correct applications of the contractual provisions to the vignettes. In all the contracts, our consumers did not, in general, understand the effect of the terms that they read. This may suggest that reading scores are not appropriate for operationalising 'intelligibility'. However, a caveat must be advanced. None of the contracts examined had a particularly low reading level. All required a post-16 educational level. Therefore, it may be the case that none of the contracts were sufficiently intelligible, and if a lower reading score were achieved then an increase in intelligibility would be detected.¹³⁹

Conclusion: are reading scores useful for assessing contractual language

Reading scores are an attractive tool for assessing contractual readability. They are easy to operate and cheap, and the process of assessment can be automated. Reading scores have the potential to function as a regulatory tool, and can be used both by businesses in drafting contracts and regulators in assessing

¹³⁶See Better Regulation Executive and National Consumer Council Warning: Too Much Information Can Harm (Final Report, November 2007).

¹³⁷M Masson and MA Waldron 'Comprehension of legal contracts by non-experts: effectiveness of plain language redrafting' (1994) 8 *Applied Cognitive Psychology* 67.

¹³⁸This is in contrast to E Greene et al 'Do people comprehend legal language in wills' (2012) 26(4) *Applied Cognitive Psychology* 500.

¹³⁹As suggested by Greene et al, *ibid*.

contract that have been drafted. However, they have both strengths and weaknesses, and it is important to bear these in mind when considering the potential utility of reading scores in assessing contractual language.

(a) The strengths of reading scores

Our empirical work suggests that reading scores can play some role in assessing plain and intelligible language, but cannot play a complete role. Our examination showed, perhaps unsurprisingly, that the general exclusions clauses of travel insurance contracts needed a high level of education to comprehend. This might suggest that the clauses contained in the general exclusions are vulnerable to fairness assessment as they are not expressed in sufficiently plain and intelligible language. Reading scores may, therefore, have a role to play in assessing contractual language.

The ease of conducting a reading scores assessment is an important factor in their favour. Reading score calculators are freely available, and computer programs can be written to enable the assessment of large numbers of contracts quickly. Evidence of readability is provided in an easily comparable format, and the results can be understood relatively easily.

However, when reading scores are used each different score formula will produce a different score for each contract. This is because of the different inputs that are taken into account. A trader may therefore be able to manipulate the reading score assigned to their contract through the choice of methodology. Therefore, if it were thought useful to use reading scores for the assessment of plain and intelligible language, it would be necessary to choose a particular reading score methodology. As has been above, this has been done in the USA, with Flesch or Flesch-Kincaid scores favoured,¹⁴⁰ particularly because these are embedded within Microsoft Word.¹⁴¹

If reading scores are to be used, a combination of scores should be used, rather than relying on a single score and a single calculator, which appears to be the case in some of the US jurisdictions. One such method is explored above, and produces the Grand Weighted Mean used in this paper. Combining scores has the benefit of smoothing some of the inconsistencies of the different methods of calculating reading scores, by taking all scores into account in making the calculation and producing a consensus as to the readability of each text under analysis. This approach is more likely to produce an operationalised measure of plain and intelligible language that assesses all dimensions of the concept.

(b) The weaknesses of reading scores

Nevertheless, further important caveats must be considered. First, as mentioned earlier, the measures that we applied to the extracts have been devised and adopted for various applications.¹⁴² While the ones used in the present analysis tend to be the most widely used, and indeed, are often the indicators used in the assessment of contracts, that is not, in and of itself, evidence that they are necessarily the best predictors of ease of text processing in the context of contracts. A typical limitation of grade-level readability estimates is that they were designed and intended for the analysis of school-age texts.¹⁴³ For this reason, they might be less sensitive to differences in complexity between highly technical texts and less appropriate for the analysis of the textual content aimed at adults to which they are routinely applied. Therefore, reading scores may be providing information about readability that does not conceptualise plain and intelligible language for adult contracting parties.

Secondly, readability measures largely rely on assumptions based on surface-level features of texts, such as the number of different words or the frequency of rare words in a text, as well as the length and

¹⁴⁰See above, text to n 107 et seq.

¹⁴¹Sirico, above n 113, at 148, and see the critique of the Microsoft Word at text to n 114 et seq.

¹⁴²Above, text to n 86.

¹⁴³See J Redish 'Readability formulas have even more limitations than Klare discusses' (2000) 24(3) ACM Journal of Computer Documentation 132 for a more in-depth analysis. See also J Selzer 'Readability is a four letter word' (1981) 18 (4) International Journal of Business Communication 23.

complexity of words and sentences. In so doing, they tend to ignore other factors such as overall text coherence,¹⁴⁴ and, perhaps more significantly, the content of the text itself.¹⁴⁵ For instance, a low-frequency or polysyllabic word will affect readability scores even though its meaning might be clearly and explicitly explained in the text, for example through a clear definition to which the reader is clearly signposted, and not pose a processing difficulty for the reader. Given this finding, reading scores appear to be better at assessing whether the contractual document is plain, rather than assessing whether it is intelligible. A document that has a low reading score may not explain the ‘potentially significant economic consequences’ of the term.¹⁴⁶ On the other hand, a term which initially appears difficult may be understandable because of the way that it is presented, taking into account factors such as definitions and layout.¹⁴⁷ Further, Pau and others, when examining the New Zealand tax code, found that changes from colons to semi-colons affected readability scores,¹⁴⁸ a change that is unlikely to affect intelligibility, and suggesting that reading scores could be manipulated without alterations to the text being made that benefit the consumer. Therefore, if reading scores are to be taken into account, they should be seen as one measure of the difficulty of the text, but this should be balanced against measures that assess whether a consumer could understand the concepts expressed by the contract.

This reflects Sirico’s critique of use of reading scores in State-level legislation in the US.¹⁴⁹ He argues that the use of reading scores in the contractual context leads to overreliance on technology, with the score providing the determination that a contract is ‘fair’ and/or ‘plain and intelligible’ (to adopt the language of the CRA 2015) if it achieves a certain readability score. He argues that ‘common sense tells us that sometimes a sentence with few words and syllables can be difficult to read and a sentence with many words and syllables can be quite comprehensible’¹⁵⁰ and reliance solely on readability scores to protect consumers can therefore fail to achieve its goal. This reflects findings that reading scores are poor predictors of information gain,¹⁵¹ meaning that they are not well equipped to assist with the determination of intelligibility.

Furthermore, readability measures computed on these bases conceptually divorce the reader’s individual characteristics from the comprehension process and paint a relatively simplified picture of what is an otherwise highly interactive activity.¹⁵² The average consumer does not feature in reading scores, and the vulnerable consumer certainly does not.¹⁵³ If the average consumer is to be the yardstick against which the language of terms is judged then simple reading scores are insufficient. It is necessary to, at least, consider the reading level of an average consumer. The Growing Up Digital Taskforce, looking at the understanding of terms and conditions by children, resolved this challenge by choosing a reading score matched to the level of education of an average 12–13 year old.¹⁵⁴ This choice is not fully explained, as the report mentions users of Instagram in the 8–11 years old range, although

¹⁴⁴SA Crossley et al ‘Text readability and intuitive simplification: a comparison of readability formulas’ (2011) 23(1) *Reading in a Foreign Language* 86.

¹⁴⁵Discussed and rejected by Benson, above n 4, at 551 et seq.

¹⁴⁶Dr Seuss *Green Eggs and Ham* (HarperCollins, 2003) has a Flesh-Kincaid grade level of -1.3, but only using single syllable words in short sentences (the only polysyllabic word in *Green Eggs and Ham* is ‘anywhere’) is unlikely to be sufficient to explain the economic effects of a consumer contract.

¹⁴⁷As alluded to by A Smith J in *Office of Fair Trading v Abbey National*, discussed above text to n 55.

¹⁴⁸C Pau et al ‘Complexity of the New Zealand tax laws: an empirical study’ (2007) 22(1) *Australian Tax Forum* 59 at 89.

¹⁴⁹Sirico, above n 113.

¹⁵⁰*Ibid.*, at 170.

¹⁵¹RG Funkhouser and N Maccoby ‘Communicating specialized science to a lay audience’ (1971) 21 *Journal of Communication* 58.

¹⁵²RG Funkhouser ‘Functional evaluation of consumer documents’ (1983) 20(3) *Journal of Business Communication* 59 at 60 argues that reading scores that are ‘developed for ease of application...’, particularly by computers, are removed from ‘the measurement of actual vocabulary difficulty’, reducing their usefulness in measuring consumer understanding.

¹⁵³For example, in cross-border transactions, where the consumer may not be a native language speaker, a reading score is unlikely to capture whether the clause would be plain and intelligible to a second language speaker.

¹⁵⁴*Growing Up Digital*, above n 56, p 8, fn 2.

Instagram's terms require users to be over 13.¹⁵⁵ This option will not be available whether the consumer is vulnerable for reasons other than age, as the expected comprehension level of the vulnerable consumer will not be so easily calculable.

Further, it is necessary to explore what a being a reasonably well-informed, reasonable observant and reasonably circumspect consumer means in the contractual context. Even if a baseline of reading comprehension could be set, it is unlikely that a simple reading score can capture whether a reasonable consumer could understand a contract. Intelligibility, and being able to calculate the economic effects of a contract, goes beyond simply being able to read a contract. Indeed, it is possible that a contract, whilst understandable to a very young audience (as demonstrated by a low reading score), would not accurately communicate the effects of a contract. In some cases some level of complexity (although not excessive levels) may contribute to understanding of economic effects, despite increasing the reading score of a particular contract.

Further, reading scores cannot account for behavioural effects. A consumer may perceive the economic effects of a contract differently depending on how the effects are framed. The consumer's understanding of the effects of the contract will vary with the presentation of the terms, despite the contracts having the same or similar reading scores. Similarly, reading scores cannot account for the effects of layout on the understanding of consumers.

It seems therefore, that reading scores cannot provide determinative evidence that consumers can understand the effects of a contract. It may be very good evidence that a consumer cannot understand the effects. A reading score cannot be the sole evidence that a contract is not in plain, intelligible language. However, it may play a role in a multifactorial approach.

(c) What role for reading scores?

The weaknesses identified above do not mean that readability has no place in the process of composing and revising a text to maximise readability and comprehension; it does suggest that they should only represent one tool in the arsenal of a regulator, trader or lawyer. In other words, while readability measures can alert the writer of a text as to its relative complexity, they cannot identify specific processing and comprehension bottlenecks. They may be able to identify whether a contract is expressed in plain language, but they cannot assess its intelligibility. A high reading score may function as good evidence that a document does not meet the plain and intelligible language requirements of the CRA 2015, but should not be the only way that one assesses compliance.


Reading scores can, when used as part of a more comprehensive evaluation of a text, play a role in assessing plain and intelligible language. In particular, they may help businesses when redrafting contracts. If reading scores are to be used, the tendency for different measures and different calculators to rely on various assumptions points to the need for an analysis approach that can at least partly offset these biases. The use of a Grand Weighted Mean, as used in this paper, is to be encouraged. However, the issues identified in this paper suggest that reading scores alone should not be relied upon by business, regulators or courts in determining that contracts are transparent, and particularly determining that they can be understood by consumers. Assessing whether a contractual clause is expressed in plain and intelligible language requires the synergistic use of other methodologies that can more directly probe the processing and understanding of texts on the part of the reader. In this sense, a typical paradigm might involve the use of methods to evaluate the ways that consumers read texts¹⁵⁶ and comprehension questions to assess understanding and retention of information.¹⁵⁷ Measures of other factors included in the CMA 2015 guidance may also be taken into account. Such an approach is

¹⁵⁵Instagram Terms of Use, clause 1, available at <https://help.instagram.com/478745558852511> (last accessed 4 November 2018). The age limit of 13 appears generally in social media terms of use (see Snapchat (<https://www.snap.com/en-GB/terms/>, clause 1); Facebook (<https://www.facebook.com/terms>, clause 4(5)); and Twitter (<https://twitter.com/tos>, clause 1)).

¹⁵⁶For example, eye-tracking. See K Conklin, A Pellicer-Sánchez and G Carrol *An Introduction to Eye-tracking: A Guide for Applied Linguistics Research* (Cambridge: Cambridge University Press, 2017).

¹⁵⁷Davis, above n 135.

unlikely to provide the simple, cheap, answers that reading scores might, but is more likely to appropriately gauge both the plainness and intelligibility of text by being able to consider both how the average consumer reads, and how a particular consumer reads.

Author ORCIDs.  Richard Hyde <http://orcid.org/0000-0002-2576-3448>; Kathy Conklin <http://orcid.org/0000-0003-2347-8018>; Fabio Parente <http://orcid.org/0000-0003-4789-9511>.

Cite this article: Conklin K, Hyde R, Parente F (2019). Assessing plain and intelligible language in the Consumer Rights Act: a role for reading scores? *Legal Studies* 39, 378–397. <https://doi.org/10.1017/lst.2018.25>