

RegData

Quantitative Analysis of State School Education Laws in India

April 2021



Copyright ©2021

First published in 2021
by Centre for Civil Society

Authors: Jayana Bedi and Prashant Narang
Centre for Civil Society

In collaboration with Mercatus Center

Data Visualisation: Ahmed Pathan

For more information and other requests, write to:

Centre for Civil Society
A-69, Hauz Khas, New Delhi – 110016
Phone: +91 11 26537456
Email: ccs@ccs.in Website: www.ccs.in

Contents

- Introduction** **1**

- Methodology** **1**
 - Word count 1
 - Binding words 1
 - How we identified binding words in state school education laws 1
 - Linguistic complexity 2

- How do states fare on the quantitative metrics?** **3**
 - Volume of state school education laws 3
 - Restrictiveness of state school education laws 3
 - Linguistic complexity of state school education laws 4
 - Flesch Reading Ease Score 4
 - Sentence length 4
 - Conditional count 5
 - Shannon entropy 5

- Conclusion** **5**

Introduction

There are two methods to assess the quality of any law: quantitative and qualitative assessment. Centre for Civil Society's (CCS) past work on school education focussed on qualitative assessment of laws. Departing from this approach, we now collaborate with the [Mercatus Center](#) to conduct a quantitative analysis of all 145 state school education laws in India.

In quantitatively analysing these laws, we calculated the following metrics: law volume (or total word count), number of restrictions a law imposes, and complexity of a law. These variables demonstrate association with economic growth and productivity (McLaughlin, Strosko, and Jones 2019). This empirical analysis will enable a new and different perspective on the education sector reforms.

Methodology

[RegData](#), an initiative of the Mercatus Center, is an effort to quantify regulation to “introduce an objective, replicable, and transparent methodology for measuring regulation.” It uses custom-made text analysis and a machine-learning algorithm to measure volume, restrictiveness, and linguistic complexity of the law.

Below we give an overview of the three metrics studied. A detailed methodology for each metric and the algorithms used can be found at the [QuantGov website](#).

Word count

Word count per state: This metric quantifies the total number of words in all laws in a state.

Average word count per state: This metric quantifies the average number of words per law in a state.

Binding words

RegData uses a text analysis program to count the number of binding words. Binding words or ‘restrictions’ in a law create an obligation to comply or limit choice sets for individual actors and include phrases such as ‘shall’, ‘must’, ‘may not’, ‘required’ and ‘prohibited’. CCS analysed a sample of 29 state school education laws to identify the binding words used.

Normalised binding words: Since binding words might be higher in laws that are lengthier, we also estimate the density of restrictiveness in each law. This metric, referred to as ‘normalised binding words’, facilitates comparison between laws and states. It highlights the average number of words after which a binding word appears. For instance, a value of 200 means that on average, a binding word appears after every 200 words in the laws of a particular state. Lower normalised binding words implies greater restrictiveness in the laws of a state.

How we identified binding words in state school education laws

CCS adopted the following 5 step approach:

Step I Preparing a database: In 2020, CCS collated a [database](#) of all 145 state laws regulating school education in India. The data set was compiled using legislations listed on all state Department websites and online sources including Laws of India, Manupatra, Bare Acts Live and Latest Laws.

Step II Sampling process: Using our database, we selected a sample of 29 laws (20% of the total laws), ensuring representation of all states. We began the sampling process by arranging all laws in chronological order and selecting every 6th law to the sample. To this list, we added laws from states with only one law on school education.

Step III Collating the binding words: All 29 laws in the sample were studied in detail to identify binding words i.e. words, appearing at least once, to restrict the actions of individuals such as those willing to set up schools, existing school owners, or non-government entities. One example of binding words is the term ‘fine’ in Section 7(8) of the Kerala Education Act, 1958. The use

of the word is restrictive in that it imposes a ‘fine’ on a school manager failing to comply with subsections 6 and 7 of Section 7 of the Act.

Step IV Preparing a restriction counter: We prepared a restriction counter for all 29 laws, recording against each law, the number of appearances of a particular binding word. This aggregate number was broken down into two categories: restrictive and nonrestrictive. The former includes only those instances when the binding word restricts actions of private actors and the latter represents all other appearances of the binding word, including ones that restrict the actions of public officials, limit their scope of power, and instill accountability.

Step V Setting a threshold to prepare the final list of binding words: For each of the 29 state school education laws in our sample, the phrases appearing in the restrictive category for more than 80 per cent of their total appearances were chosen as binding words. The rest were discarded, as were those that appeared in the state law only once. The following 19 binding words made it to the final list: ‘bound’, ‘binding’, ‘deemed to be guilty’, ‘comply’, ‘impose’, ‘shall be punishable’, ‘punished’, ‘fine’, ‘imprisonment’, ‘withdraw’, ‘withdrawal’, ‘suspend’, ‘suspended’, ‘supersede’, ‘shall be liable’, ‘discontinue’, ‘contravene’, ‘contravenes’, ‘seize’.

Linguistic complexity

This metric measures the complexity of a given law based on how it fares on the four sub-metrics conceptualised in RegData (described below). These sub-metrics provide an estimate of how easy or difficult it is to comprehend a law. A law that is tough to comprehend may also increase the compliance costs for regulated entities in terms of effort, time and money (McLaughlin et al. 2020).

1. *Shannon entropy:* The complexity of a text is directly proportional to its ‘Shannon entropy’ score. This score signifies the likelihood of encountering new words and concepts in a given body of law. These scores are best understood in relative terms. See [here](#) for more information.
2. *Sentence length:* This metric signifies the average length of sentences in a law. Longer sentences tend to be more difficult to read.
3. *Conditional count:* Terms like ‘if’, ‘but’ and ‘provided’ create logical branches in a legislation and are called ‘branching words’. Conditional count measures their total number in a legislation.
4. *Flesch Reading Ease score:* This metric measures the ease of reading of a legislation. Based on the average sentence length and syllables per word, each legislation is assigned a score on a scale of 0 to 100, 0 being the most difficult to read and 100 being the easiest to read. The higher a legislation’s Flesch score, the easier it is to read. A Flesch reading score may be negative as well.

How do states fare on the quantitative metrics?

The metrics discussed above signal the extent of regulatory burden and restrictions on schools. They also indicate the ease with which regulatees can understand the law. Our [interactive dashboards](#) provide details of all three quantitative metrics across all state school education laws. Below we highlight some of our key findings:

Volume of state school education laws

Word count by law: The average number of words in a state school education law is 4,700 words. The total word count of 145 state school education laws is over 6,80,000 words. Arunachal Pradesh has the highest average word count of laws.

Word count by state: The average word count per state is over 22,000 words. West Bengal has the most voluminous laws, with a total word count of 61,458. The other states with voluminous laws are Telangana (43,333), Karnataka (43,267), Maharashtra (42,811) and Uttar Pradesh (41,651). States with the lowest total word count include Nagaland (1,002), Chhattisgarh (4,322), Orissa (6,056), Sikkim (7,006) and Kerala (7,312).

Name of the law	Word count
West Bengal Primary Education Act, 1973	35,845
Karnataka Education Act, 1983	32,212
Arunachal Pradesh Education Act, 2010	27,221
Telangana Education Act, 1982	25,815
The Andhra Pradesh Education Act, 1982	24,261
Uttaranchal School Education Act, 2006	14,711
West Bengal Board Of Secondary Education Act, 1963	13,681
Gujarat Secondary and Higher Secondary Education Act, 1972	13,530
Maharashtra Secondary and Higher Secondary Education Boards Act, 1965	12,644
Puducherry Board of Secondary and Higher Secondary Education Act, 2003	11,389

On average these 10 laws have 21,131 words.

Restrictiveness of state school education laws

On average, a state school education law uses 10 binding words. Arunachal Pradesh has the most restrictive school education laws with the lowest score of normalised binding words (231) i.e. on average, a restrictive term appears after every 231 words in a law in Arunachal Pradesh. This is followed by Chhattisgarh and Karnataka, at 254 and 259 words respectively (Table 2). Nagaland is the only state that uses binding words only once in the law.

Name of the State	Normalised binding words
Arunachal Pradesh	231
Chhattisgarh	254
Karnataka	259
Tamil Nadu	292
Delhi	312

The most commonly used restrictive terms are ‘fine’ (229 times) and ‘comply’ (197 times). Karnataka Education Act, 1983 has the highest usage of the term ‘fine’ (31 times). Arunachal Pradesh has the highest usage of the term ‘comply’ (14 times).

The 5 least restrictive states and union territories include Orissa, Nagaland, Jharkhand, Jammu and Kashmir, and West Bengal.

Linguistic complexity of state school education laws

The complexity of state school education laws can be determined based on 4 sub-categories: Flesch Reading Ease Score, Sentence length, Conditional count, and Shannon Entropy.

Flesch Reading Ease Score

The average Flesch score of state school education laws is 30. The band 30-50 signifies that an individual must be at least a college graduate to understand the law.

Mizoram is the worst performing state on this metric, with an average Flesch score of 2.5, followed by Delhi (7.5). The three worst performing laws, with a negative Flesch score, are Mizoram Education Act, 2003; Maharashtra Educational Institutions (Transfer of Management) Act, 1971; and U.P. Self-Financed Independent Schools (Fee Regulation) Act, 2018. A negative score implies that even a college graduate will struggle to read and understand the law.

The three best performing states on this metric are Manipur (49.75), Telangana (46) and Himachal Pradesh (44). However, even their score falls in the ‘difficult to read’ category and can only be understood by a college graduate. The three easiest to read laws are U.P. Educational Institutions (Taking-Over of Management) Act, 1976 (76); West Bengal Council Of Higher Secondary Education Act, 1975 (65); and Manipur Elementary and Secondary Education Act, 1972 (64).

Name of the State	Flesch Reading Ease Score
Mizoram	2.5
Delhi	7.5
Puducherry	9
Chhattisgarh	10
Goa	14.25

Sentence length

The average sentence length across all 145 laws is 39 words. Long sentences add to the complexity of a legislation.

Three states and union territories with highest sentence length (on average) include Chhattisgarh (68), Delhi (66) and Puducherry (60.83). The worst performing laws on this metric include West Bengal Non-Government Educational Institutions and Local Authorities (Control of Provident Fund of Employees) Act, 1983 (81.52); Mizoram Education Act, 2003 (74.55); and Orissa Aided Educational Institutions (Appointment of Hindi Teachers Validation) Act, 1992 (71.5).

Punjab is the best performing state on this metric, with an average sentence length of 10 words. This is followed by Telangana (22) and Manipur (24.78). Three best performing laws include: U.P. Educational Institutions (Taking-Over of Management) Act, 1976; Telangana Private Educational Institutions Maintenance Grant (Regulation) Act, 1995; and West Bengal Council Of Higher Secondary Education Act, 1975.

Conditional count

On average, state school education laws have 30 branching words. Arunachal Pradesh Education Act, 2010 has the highest conditional count (151). The Nagaland Board of Secondary and Higher Secondary Education Act, 1973 uses only 4 branching words. This is followed by Orissa with approximately 5 branching words used per law.

Shannon entropy

The average Shannon entropy score of state school education laws is 8.23. The three worst performing states on this metric are Arunachal Pradesh (9.58), Uttarakhand (9.15) and Manipur (8.92). The three best performing states on this metric are Nagaland (7.44), Odisha (7.49) and Tamil Nadu (7.85).

Finally, we find a strong positive correlation between the volume of a law and its complexity. As the volume of a law increases, its Shannon entropy score (or complexity) also increases.

Conclusion

The quantitative metrics highlighted above indicate the extent of regulatory burden on schools across the country. Our analysis helps compare states to one another and identify areas of improvement. However, without a baseline, it is difficult to comment on the level of restriction, volume or complexity best suited for education laws.

While we did not score states based on these quantitative metrics, some states fare well on most metrics. Nagaland and Odisha are the best performing states on law volume, restrictiveness and Shannon entropy. Arunachal Pradesh is the worst performing state on all three metrics. On reading ease, laws in Manipur and Telangana fare well. Manipur's performance on linguistic complexity is mixed: while it is the best performing state on reading ease, it is amongst the bottom three states on Shannon entropy.

References

- McLaughlin, Patrick A, Kofi Ampaabeng, Ethan Greist, Jonathan Nelson, Thurston Powers, and Stephen Strosko. 2020. “RegData 3.2 User’s Guide.” *Mercatus Center, George Mason University*. https://new-website-files.s3.amazonaws.com/regdata3_2_users_guide.pdf.
- McLaughlin, Patrick A, Stephen Strosko, and Laura Jones. 2019. “RegData Canada: A Snapshot of Regulatory Restrictions in Canada’s Provinces.” *Mercatus Center, George Mason University, Regulatory Snapshot*.