Journal of Education
1–13

DOI: 10.1177/00220574211002199
journals.sagepub.com/home/jex
SAGE

Classification of Action Verbs of Bloom's Taxonomy Cognitive Domain: An Empirical Study

Syaamantak Das[1], Shyamal Kumar Das Mandal[1], and Anupam Basu[1]

Abstract

It was observed from previous researches that Bloom's Taxonomy action verbs (BTAVs) are overlapping in multiple cognitive levels, causing ambiguity about the real sense of the word. A data set of BTAVs was ranked using a statistical classification. Four categories of BTAVs were identified, out of which 153 BTAVs (86.44%) are classified into a single cognitive level. In the remaining 24 BTAVs, 21 follow a specific transition from one cognitive level to another. The proposed research reduced the cognitive level's ambiguity and made the classification more robust, making it easier for teachers to create and set the cognitive levels.

Keywords

cognitive science, Bloom's Taxonomy action verbs, question analysis

Introduction

Any form of learning depends on three significant pivots—(a) learning goals, (b) teaching–learning activities, and (c) assessment and feedback. This tri-dimension feature is mentioned in several previous literature pieces, especially in Fink's (2013) backward design model. The model focuses on the "importance of aligning learning goals with learning activities and assessments." The learning goals are also known as learning objectives, instructional objectives, or learning outcomes (LOs). LOs refer to statements that state the knowledge and skills that are going to be gained by the learner after successful completion of the course lesson (Allan, 1996; Otter, 1995). The teaching–learning activities refer to the usage of teaching–learning materials (TLM), and an assessment relates to evaluation methods. Textbooks form a significant part of TLM (Sheldon, 1988; Stelzer et al., 2009), mainly school textbooks, as they provide both content and assessment questions.

Bloom's Taxonomy (Bloom, 1956) is a widely used methodology to describe LOs in educational documents such as setting up curriculum, performing the objectives-based evaluation on learner's achievement, and aligning curriculum and assessment (Lee et al., 2017). Bloom's Taxonomy describes six levels of cognitive development—"Knowledge (Remember), Comprehension (Understand), Application, Analysis, Evaluation, Synthesis (Create)," ranging from simple remembering to complex and critical reasoning abilities (DeWaelsche, 2015; Swart & Daneti, 2019). *Knowledge* and *Comprehension* were more concrete and straightforward than *Synthesis* and *Evaluation*, which were more complex and abstract (Krathwohl & Anderson, 2010).

The original taxonomy of 1956 was based on action verbs in noun form to identify the cognitive levels. Krathwohl and Anderson (2010) observed that as a taxonomy of cognitive levels, the noun forms used for the original cognitive category names were inappropriate. They observed that the cognitive level *Create* is more than making new knowledge fit with existing knowledge as cognitive level *Synthesis* suggests. Later in 2002, the revised taxonomy by Krathwohl (2002) focused on using the verb form of the words (DeWaelsche, 2015).

However, one drawback of both the old and revised versions was that some of these action verbs kept overlapping in multiple taxonomy levels, causing an ambiguity about the real sense of the cognitive level. Several other research across the world created similar data sets based on these words, which added uncertainty.

Claudia Stanny's work (Stanny, 2016) provided a significant compilation of these action verbs. The paper compiled 30 different Bloom's Taxonomy action verbs (BTAVs) data set and prepared a list of 176[1] action verbs. The article raised an important question—Can unambiguous recommendations

[1]Indian Institute of Technology Kharagpur, India

Corresponding Author:
Syaamantak Das, Centre for Educational Technology, Indian Institute of Technology Kharagpur, Kharagpur 721302, India.
Email: syaamantak.das@gmail.com

be made about which BTAVs correspond to specific levels of cognitive skills by the structure of Bloom's Taxonomy? Our paper attempted to answer that question.

Our Contributions

The proposed research paper uses the existing data set (Stanny, 2016) and performs a statistical ranking of the 177[1] words across the 30 data sets and classifies them into four levels. The first level consists of unique words and has an occurrence frequency above the set threshold value. The next level consists of unique words but has an occurrence frequency lower than the threshold value. The third level consists of overlapping words in two or more specific levels (transitional) and has occurrence frequency above the set threshold value. The fourth and final level consists of words that occur in multiple levels (more than 50% of the six levels) and has a frequency more than the threshold value.

To validate the classification, a set of 185 questions containing BTAVs were taken from the National Council of Educational Research and Training (NCERT) data set of Rakesh Agrawal's work in Microsoft Search Lab (Agrawal et al., 2012, 2014). Five human annotators manually annotated the 185 questions. The action verbs' distribution across the cognitive domain was observed and matched with the proposed classification. Previous research literature that performed similar experiments was used for result comparison. The findings show that both the previous result and the proposed model's result have a similar verbs distribution over the cognitive level. Thus, the validation of the proposed classification model was proven.

Significance of the Study

One of the essential Learning Objective goals is to describe the specific skills and knowledge that the student will acquire. Bloom's Taxonomy is a standard approach used by educational institutes globally to write measurable Learning Objectives and describe them. The other goals that also get influenced by Bloom's Taxonomy are aligning the learning materials and assessments as per the curriculum. All these goals are dependent on the verbs that the instructors select, when they describe the thinking skills required for a given LO. Fink (2013) proposed aligning learning objectives with learning materials and assessment, based on his backward design model. The idea was that a teacher first creates a course with Learning Objectives. Then the teacher identifies the learning materials and makes assignments so that the learners can acquire and practice the skills. Furthermore, the learners are tested on assessments that check whether the intended LO is being satisfied or not. Although several other taxonomies are available, like SOLO (Structure of the Observed Learning Outcome)—which helps to frame better content—Bloom's Taxonomy appears to be one of the time-tested models applied in all three learning goals.

However, two challenges are there to answer—(a) Can there be a consensus about the level of thinking skill specified by a verb? (b) Can there be unambiguous recommendations about verbs corresponding to specific cognitive skills levels based on Bloom's Taxonomy structure? The idea is to categorize verbs that are consistently aligning with particular levels of Bloom's Taxonomy. This scaling down will reduce the ambiguity, which is existing in the present structure. A verb-based framework's major problem is that the language is notoriously flexible (Stanny, 2016). A verb can be associated with low and high Bloom's Taxonomy levels based on the word's context. Thus, it becomes essential to understand the context in which the verb is placed. In the context of a specific question or learning material, the taxonomy's action verb has specific and unambiguous meaning. If taken out of context, the verbs no longer remain unambiguous.

Literature Review

Distribution of Bloom's Taxonomy Cognitive Levels Over School Curriculum

For sampling purposes, the proposed research work considered recent related works (2010–2015) on the distribution of Bloom's Taxonomy's cognitive levels over school education curriculum. Some recent research works (Abdelrahman, 2014; Assaly & Smadi, 2015; Igbaria, 2013; Lee et al., 2015; Riazi & Mosalanejad, 2010) have considered Bloom's Taxonomy as a benchmark for cognitive level assessment of school-level curriculum, which consists of Learning Objectives, learning materials, and assessment questions. The study by Riazi and Mosalanejad (2010) on English textbooks in Iranian senior high schools and pre-university colleges indicates that lower-order cognitive skills were more prevalent in all grades than higher-order ones. The lower-order cognitive levels accounted for 75.3%, whereas higher-order skills accounted for 24.7%. The work by Lee et al. (2015) in primary science curricula from Korea and Singapore using revised Bloom's Taxonomy showed that 33.3% of cognitive items fall into *Remember/Knowledge* level, whereas 54% falls into *Understand/Comprehension* level. Similar trends were also observed in the analysis of questions. The research by Assaly and Smadi (2015) showed that *Knowledge* and *Comprehension* cognitive levels accounted for about 57% of the questions. The research work of Igbaria (2013) analyzed WH questions. The results showed that *Knowledge* and *Comprehension* accounted for 23.09% and 29.66%, which is again more than half of the entire distribution. The work by Abdelrahman (2014) analyzed 10th-grade English language textbooks and observed that most of the questions were within the first two levels,

that is, *Remembering* and *Understanding* (55.11%). Thus, it can be concluded that it is a global trend to focus more on low-level cognitive skills as far as school education is concerned.

Bloom's Taxonomy and LO

As observed in the work of Swart and Daneti (2019), the results show that the first and second cognitive levels of Bloom's Taxonomy (*Knowledge* and *Comprehension [Understanding]*) constitute about 58% of the given LOs. The third cognitive level (*Application*) is about 27%. The highest two levels (*Synthesis [Create]*, and *Evaluation*) constitutes 15%. This observation was based on the LOs of an electronics course module (termed Electronics 1) offered in two universities of Africa and Europe.

Bloom's Taxonomy and Assessment Questions

Jones et al. (2009) indicate that academics in electronics use a greater amount of lower-level cognitive questions than higher cognitive level questions in their assessments. The data were collected on the examination papers for one academic year to "second and final year students on eight programs." Each question was "evaluated and categorized into low, intermediate, or higher-order cognitive questions" based on the verb list provided by Dalton and Smith (1989).

Bloom's Taxonomy Criticisms

As stated by Lee et al. (2017), that despite the practicality and simplicity of the model, Bloom's Taxonomy is criticized for generalized and uni-dimensional "domains of knowledge and skills that could not clearly explain the levels." Moreover, the levels of cognitive demands in the analysis of instructional objectives for students' learning and assessment plans also remain ambiguous. Although the revised Bloom's Taxonomy tries to overcome the generalization of cognitive dimensions, "there is still the challenge of identifying the level of thinking."

Data Set Details

BTAV Corpus

The data set for BTAVs was obtained from the previous reference work of Stanny (2016). The data set's BTAVs were collected from 30 websites, which include academic centers such as Cornell, UNCC, UCI, Kent, Carnegie Mellon, UniAlberta, and other educational institutes. WH words ("how," "what," "when," "where," "which," "who," "why") were not considered as a BTAV. Words were normalized to their base form (lemma; e.g., compared became compare). Occurrence frequency was calculated based on the number of times a particular list assigned a verb to a given level of Bloom's Taxonomy. This value ranged from 1 (only one list assigned the verb with that category) to 30 (every list assigned the verb with that category). The paper used two threshold values—(a) Four (lenient criterion), that is, at least four or more lists contained the verb, and (b) 10 (stringent criterion), that is, at least 10 or more lists contained the verb. Verbs also duplicated across the category with values ranging from 1 (verbs assigned to a single group) to 6 (verbs assigned to every group). Based on the lenient criterion, the verb corpus had a collection of 177 distinct verbs. Out of 177 verbs, 118 (67%) verbs appeared in one cognitive level. The remaining 59 verbs had a distribution of 38 (21.4%) in two cognitive levels, 14 (7.9%) in three cognitive levels, five (2.8%) in four cognitive levels, and two (1.1%) in five cognitive levels. In the stringent criterion, the collection was reduced to 104 verbs. Eighty-three verbs (79.8%) were in one category. The rest 21 words (20.2%) appear in more than one category—18 words (17.3%) in two cognitive levels and three words (2.9%) in three cognitive levels. The paper observed that even after taking this stricter (stringent) criterion, the reduced corpus of verbs provided "evidence of multiple interpretations of skill described by a verb." The cognitive distribution of both stringent and lenient criteria is shown in Table 1. The action verbs of the lenient criterion are shown in Table 2.

Table 1. Cognitive Level Distribution of Stanny's (2016) Data Set.

Criteria type	K	C	AP	AN	E	CR	Total
Stringent (≥10)	20	25	23	22	18	20	128
Lenient (≥4)	34	46	44	48	43	48	263

Note. K = Knowledge; C = Comprehension; AP = Application; AN = Analysis; E = Evaluation; CR = Create.

For a cross-reference, another data set of BTAVs by Center for Advancing Teaching and Learning through Research, Northeastern University[2] was used, which was not present in the previous compilation by Stanny. The data set consists of 226 action verbs, out of which 167 are distinct. When these 167 distinct data were compared with the previous data set of Stanny's, 35 new unique words were identified. These words are shown in Table 3. As no occurrence frequency is available, these words' inclusion in these levels is assumed to be accurate, considering the last data set was prepared in 2018. The distribution of additional new 35 unique words over cognitive level is shown in Table 3.

BTAV Question Corpus

The question corpus was obtained from the NCERT textbook data set. The NCERT textbook data set for this proposed research was a 10th-grade school textbook of Economics,

Table 2. Bloom's Taxonomy Action Verbs of Stanny's (2016) Lenient Criterion With Occurrence Frequency Denoted by *f*.

Knowledge	*f*	Comp.	*f*	Appl.	*f*	Analysis	*f*	Eval.	*f*	Create	*f*
Arrange	6	articulate	4	act	19	analyze	24	appraise	22	arrange	22
Choose	4	associate	4	adapt	4	appraise	11	argue	12	assemble	14
Cite	17	characterize	4	apply	22	break	8	arrange	5	categorize	7
Copy	4	cite	4	back/back up	5	break down	7	assess	17	choose	7
Define	21	clarify	5	calculate	10	calculate	9	attach	4	collect	9
Describe	14	classify	18	change	9	categorize	19	choose	10	combine	14
Draw	5	compare	11	choose	11	classify	10	compare	18	compile	7
Duplicate	7	contrast	7	classify	6	compare	24	conclude	13	compose	19
Identify	20	convert	13	complete	5	conclude	6	contrast	8	construct	29
Indicate	4	defend	12	compute	10	contrast	19	core	6	create	19
Label	21	demonstrate	6	construct	13	correlate	5	counsel	4	design	24
List	27	describe	22	demonstrate	20	criticize	11	create	4	develop	18
Locate	10	differentiate	8	develop	4	debate	8	criticize	11	devise	13
Match	14	discuss	21	discover	8	deduce	6	critique	14	estimate	5
Memorize	10	distinguish	12	dramatize	16	detect	7	decide	4	evaluate	4
Name	22	estimate	11	employ	16	diagnose	4	defend	15	explain	8
Order	5	explain	28	experiment	6	diagram	12	describe	4	facilitate	4
Outline	11	express	17	explain	5	differentiate	20	design	4	formulate	18
Quote	7	extend	11	generalize	5	discover	4	determine	6	generalize	7
Read	4	extrapolate	5	identify	4	discriminate	11	discriminate	9	generate	11
Recall	24	generalize	11	illustrate	18	dissect	6	estimate	15	hypothesize	8
Recite	12	give	4	implement	4	distinguish	21	evaluate	16	improve	5
Recognize	14	give examples	8	interpret	15	divide	12	explain	9	integrate	4
Record	13	identify	14	interview	6	evaluate	4	grade	4	invent	10
Relate	11	illustrate	9	manipulate	10	examine	18	invent	8	make	6
Repeat	20	indicate	8	modify	12	experiment	9	judge	25	manage	8
Reproduce	11	infer	15	operate	17	figure	4	manage	15	modify	10
Review	4	interpolate	5	organize	4	group	4	mediate	9	organize	21
Select	16	interpret	17	paint	4	identify	7	prepare	12	originate	9
State	23	locate	10	practice	15	illustrate	8	probe	4	plan	21
Tabulate	4	match	7	predict	9	infer	14	rate	5	predict	8
Tell	4	observe	5	prepare	11	inspect	8	rearrange	19	prepare	12
Underline	7	organize	5	produce	13	inventory	9	reconcile	12	produce	13
Write	5	paraphrase	22	relate	12	investigate	7	release	6	propose	9
		predict	12	schedule	11	order	5	rewrite	4	rate	21
		recognize	11	select	4	organize	6	select	5	rearrange	8
		relate	7	show	13	outline	10	set up	15	reconstruct	9
		report	10	simulate	5	point out	12	supervise	9	relate	8
		represent	4	sketch	17	predict	4	synthesize	16	reorganize	9
		restate	15	solve	19	prioritize	4	test	8	revise	12
		review	15	translate	5	question	12	value	7	rewrite	7
		rewrite	12	use	25	relate	17	verify	9	role-play	4
		select	7	utilize	4	select	12	weigh	5	set up	9
		summarize	20	write	5	separate	10			specify	5
		tell	7			solve	8			summarize	7
		translate	21			subdivide	10			synthesize	4
						test	14			tell/tell why	5
										write	17

Geography, and Science (Biology). The questions were collected from both in between the chapters and the end of the chapters. There were 434 questions altogether, out of which 184 (42%) had BTAVs. The number of BTAVs question for Economics, Geography, and Science (Biology) are 101, 52, and 31, respectively.

Table 3. The Cognitive Level Distribution of Additional New 35 Unique Words From Northeastern University Data Set.

Knowledge	Comprehension	Application	Analysis	Evaluation	Create
enumerate	Ask	Administer	advertise	Consider	Anticipate
omit	Research	Chart	connect	Convince	collaborate
retell	Trace	Establish	focus	Editorialize	intervene
visualize	Transform	Teach		find errors	negotiate
		Transfer		Justify	speculate
				Measure	structure
				Persuade	validate
				Rank	
				Recommend	
				Reframe	
				Score	
				Support	

Cognitive Level Annotation for Question Corpus

The questions were annotated by five independent annotators who are trainee graduate teachers. The number of annotators was based on Bayerl's work (Bayerl & Paul, 2011), which provides two critical elements for selecting annotators. First, annotators should be domain experts. Second, the number of annotators should lie between three to five, depending on how crucial the annotation task is. The inter-annotator agreement was calculated using the Fleiss kappa (Hallgren, 2012). For Fleiss's kappa, only one category per example and an annotator is considered. The category chosen was the one that the annotators agreed on. According to Landis and Koch (1977), kappa values between 0.01 and 0.20 represent slight agreement, those between 0.21 and 0.40 represent fair agreement, those between 0.41 and 0.60 represent moderate agreement, and those greater than 0.60 represent the substantial agreement. For measuring the percentage of agreement, a rule of thumb is that an agreement of at least 70% is acceptable as seen in the work of Stemler (2004). The kappa value for Bloom's Taxonomy ranges between 0.57 $\pm$ 0.04 and 0.73 $\pm$ 0.04, as observed in the work of Plack et al. (2007). The inter-annotator agreement for the questions of Economics, Geography, and Science (Biology) is 0.74, 0.60, and 0.61, which are substantial agreement, moderate agreement, and substantial agreement, respectively. When compared with previous similar works (Swart & Daneti, 2019), the distribution of cognitive level was found to be similar (Figure 1). The comparison is shown in Table 4.

The Methodology of the Classification Algorithm

It was observed from the data set that there could be two parameters for classification. First is frequency count, and second is whether the word belongs to a single level or multiple cognitive levels. The lenient threshold value (occurrence frequency $\geq$4) was considered for frequency count. And stringent criterion threshold value (occurrence frequency $\geq$10, out of 30 data sets) was considered for classification count of action verbs. The generic computation algorithm for classification is shown in Algorithm 1.

Four possible classification levels were formed based on these parameters, which are as follows.

Level 1—Unambiguous

The first level consists of words having a higher occurrence frequency above the set threshold value in either a single cognitive level or multiple cognitive levels. If the same word belongs to multiple cognitive levels and has a lower occurrence frequency than the threshold, it is removed from that level, for example, "define," "cite." "Define" occurs only in *Knowledge* level and has the occurrence frequency of 21. "Cite" occurs in *Knowledge* level with occurrence frequency of 17 as well as in *Comprehension (understand)* level with occurrence frequency of 4. Thus, "cite" is considered for *Knowledge* level.

It was observed that there exist 83 unique BTAVs with an occurrence frequency of more than 10. These 83 BTAVs are considered for Level 1. The distribution of these 83 BTAVs is shown in Table 5.

Level 2—Unambiguous With a Lower Threshold Value

For the second level, words having a lower occurrence frequency than the set threshold value in either a single cognitive level or multiple cognitive levels. In the case of a single cognitive level, the word is assigned at that particular level. In the case of multiple cognitive levels, the word is assigned to whichever cognitive level has greater occurrence frequency, for example, "duplicate," "indicate." "Indicate"

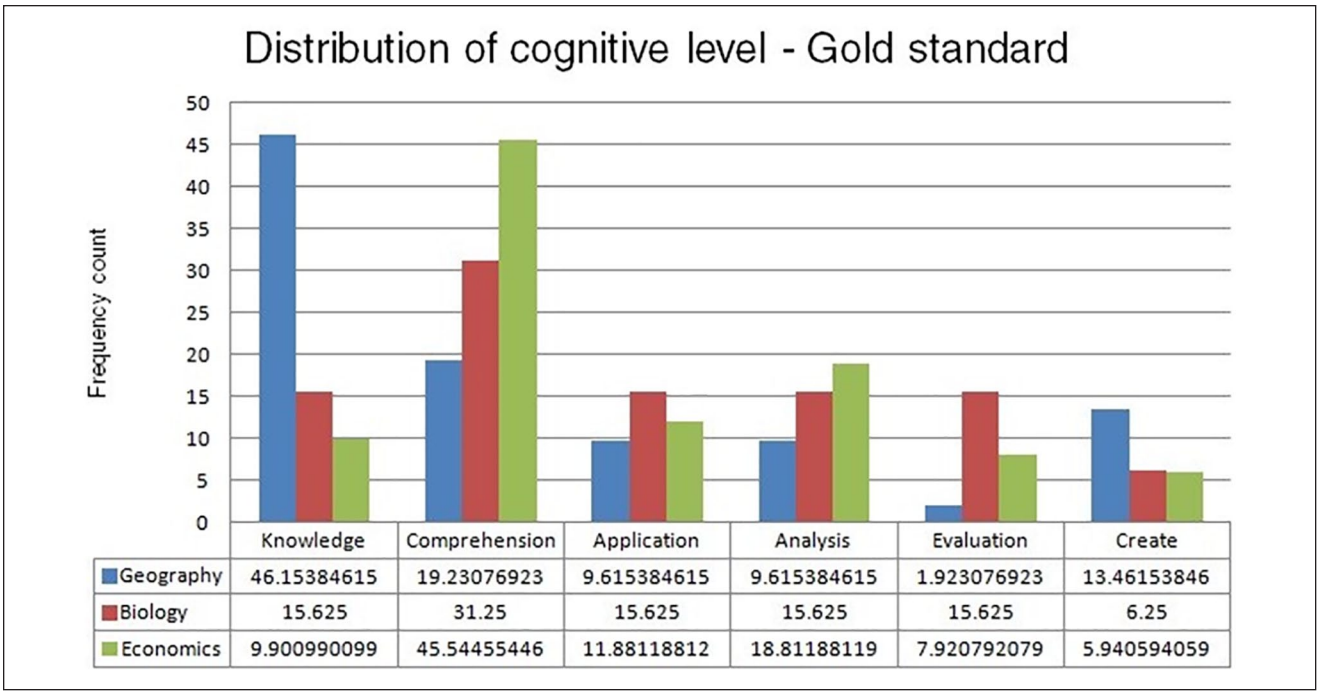

Figure 1. Distribution of cognitive levels of question corpus (gold standard).

Table 4. Comparison of the Distribution of Cognitive Level.

	K-C (%)	AP-AN (%)	E-CR (%)
Swart's distribution (Swart & Daneti, 2019)	58	27	15
Proposed work's distribution	55.90	27	17

Note. K = Knowledge; C = Comprehension; AP = Application; AN = Analysis; E = Evaluation; CR = Create.

occurs in Comprehension (understand) level and Knowledge level with an occurrence frequency of 8 and 4, respectively. Thus, "indicate" is assigned to Comprehension (understand) level. In the case of the word "duplicate," it belongs only to Knowledge level with occurrence frequency of 7.

It was observed that there are 70 unique BTAVs with an occurrence frequency of less than 10. These 70 BTAVs are considered for Level 2. The distribution of these 70 BTAVs is shown in Table 6.

Level 3—Transitional Verbs

The third level consists of words having a higher occurrence frequency above the set threshold value in multiple cognitive levels. If the same word has a lower occurrence frequency than the threshold value in other levels, it is not considered. It was observed that these words have a specific pattern, that is, it indicates a transition from one cognitive level to another. This transition can be single level (transition into the immediate next cognitive level) or multi-level (transition into multiple cognitive levels). For example, "interpret" (single-level), "choose" (multi-level). "Interpret" occurs both in *Comprehension (understand)* level and *Application* level, and has the occurrence frequency of 17 and 15, respectively. "Choose" occurs in four levels—*Knowledge*, *Application*, *Evaluation*, *Create* with an occurrence frequency of 4, 11, 10, and 7, respectively. Thus, "choose" is considered for *Application* and *Evaluation* level, and it is a multi-level transitional verb.

Transitional BTAVs can be further categorized into single-level transition and multi-level transition. Single-level transition means that the action verbs show a transfer of learning from one level to its immediate next level. Examples of such transitions are "Describe," "Identify," "locate," "recognize" which exists in both *Knowledge* and in *Comprehension* level and has an occurrence frequency of more than 10. An exceptional case of a small group of three words—"discover," "experiment," and "order"—with less than threshold frequencies is also a part of the single- and multi-level transition. When used to identify the cognitive level, each of these words makes the learner transfer the cognition from one level to another. For example, unless the learner

Algorithm 1 The classification algorithm.

Input: List of action verbs W, where W_f is word with occurrence frequency f of W in C_i, where C is Cognitive level, $i = 1$ to 6, Threshold $T \geq 10$
Output: $W \longrightarrow L_i$, where L_i is classification level, and $i = 1$ to 4

```
for each word W ∈ C_i do
#Category : Unambiguous 1
    if (W_f(C_i) ≥ T ∧W∉ C_j ) then
        put W in L_1
    else
        if (W_f(C_i) ≥ T ∧W∈ C_j, ∧W_f(C_j)<T) then
            put W in L_1
#Category : Unambiguous 2
        else
            if (W_f(C_i)<T ∧W∉ C_j,)  then
                put W in L_2
            else
                if (W_f(C_i)<T ∧W∈ C_j, ∧W_f(C_i) >W_f(C_j))  then
                    put W in L_2
#Category : Transitive
                else
                    if (W_f(C_i)∧W∈ C_j,∧W_f(C_j) >T) then
                        put W in L_3
#Category : Ambiguous
                    else
                        if (W_f(C_i) ≥ T ∧W∈ C_j ∧W_f(C_j) ≥ T ∧W∈ C_k
∧W_f(C_k) ≥ T) then
                            put W in L_4
                        end if
                    end if
```

Algorithm 1. The classification algorithm.

has knowledge about a topic, he or she will not be able to *understand* (*Comprehension*) it. Similar such transitions are observed in words like "interpret," which occurs in *Comprehension* and *Application*; "appraise," and "criticize," which occurs in *Analysis* and *Evaluation*. The word "discover," and "experiment" although has lower threshold value, belongs to both *Application* and *Analysis*. A list of single-level transitional verbs is shown in Table 7.

Table 5. Unique BTAVs With 10 or More Occurrence Frequency.

Knowledge	*f*	Comprehension	*f*	Application	*f*	Analysis	*f*	Evaluation	*f*	Create	*f*
Cite	*17*	convert	*13*	act	*19*	Analyze	*24*	argue	*12*	Arrange	*22*
Define	*21*	discuss	*21*	apply	*22*	categorize	*19*	assess	*17*	assemble	*14*
Label	*21*	explain	*28*	calculate	*10*	Contrast	*19*	conclude	*13*	combine	*14*
List	*27*	express	*17*	compute	*10*	Diagram	*12*	critique	*14*	compose	*19*
Match	*14*	extend	*17*	demonstrate	*20*	differentiate	*20*	evaluate	*16*	create	*19*
memorize	*10*	generalize	*11*	dramatize	*16*	discriminate	*11*	judge	*25*	design	*24*
Name	*22*	paraphrase	*22*	employ	*16*	Divide	*12*	manage	*15*	develop	*18*
Recall	*24*	predict	*12*	illustrate	*18*	Examine	*18*	rearrange	*19*	devise	*13*
Recite	*12*	report	*10*	manipulate	*10*	point out	*12*	reconcile	*12*	formulate	*18*
Record	*13*	restate	*15*	operate	*17*	Question	*12*	set up	*15*	generate	*11*
Repeat	*20*	review	*15*	practice	*15*	Separate	*10*	synthesize	*16*	invent	*10*
reproduce	*11*	rewrite	*12*	schedule	*11*	subdivide	*10*			organize	*21*
State	*23*	summarize	*20*	show	*13*	Test	*14*			plan	*21*
		translate	*21*	sketch	*17*					rate	*21*
				solve	*19*					revise	*12*
				use	*25*					write	*17*

Note. BTAVs = Bloom's Taxonomy action verbs.

Table 6. Unique BTAVs With Less Than 10 Occurrence Frequency.

Knowledge	*f*	Comprehension	*f*	Application	*f*	Analysis	*f*	Evaluation	*f*	Create	*f*
duplicate	*7*	associate	*4*	change	*9*	breakdown	*7*	decide	*4*	Compile	*7*
quote	*7*	characterize	*4*	complete	*5*	correlate	*5*	grade	*4*	facilitate	*4*
read	*4*	give examples	*8*	backup	*5*	deduce	*6*	weigh	*5*	hypothesize	*8*
tabulate	*4*	indicate	*8*	implement	*4*	dissect	*6*	counsel	*4*	integrate	*4*
copy	*4*	represent	*4*	interview	*6*	prioritize	*4*	mediate	*9*	originate	*9*
draw	*5*	clarify	*5*	paint	*4*	survey	*7*	probe	*4*	Propose	*9*
underline	*7*	extrapolate	*5*	utilize	*4*	break	*8*	release	*6*	role-play	*4*
		Give	*4*	adapt	*4*	detect	*7*	supervise	*9*	Improve	*5*
		interpolate	*5*	simulate	*5*	diagnose	*4*	verify	*9*	Make	*6*
		articulate	*4*			figure	*4*	attach	*4*	Specify	*5*
		observe	*5*			inspect	*8*	core	*6*	tell / tell why	*5*
		Tell	*7*			inventory	*9*	determine	*6*	Collect	*9*
						investigate	*7*	value	*7*	reconstruct	*5*
						debate	*8*			reorganize	*9*
						group	*4*				

Note. BTAVs = Bloom's Taxonomy action verbs.

In multi-level transitional verbs, it was observed that there is a cognitive level jump, that is, an intermediate cognitive level has little or no influence. Examples of such multi-level transitional verbs are "outline" and "select" (occurring in *Knowledge* and *Analysis*); "distinguish," "infer," and "classify" (occurring in *Comprehension* and *Analysis*); "defend" and "estimate" (occurring in *Comprehension* and *Evaluation*); "choose" (occurring in *Application* and *Evaluation*); "construct," "modify," and "produce" (occurring in *Application* and *Create*).

Table 8 shows the list of multi-level transitional verbs. The verb *order* belongs to *Knowledge* and *Analysis*, that is, multi-level transition, whereas others belong to single-level transition.

Level 4—Ambiguous

For the fourth level, words have a higher occurrence frequency than the set threshold value in multiple groups. The set of words are genuinely ambiguous and are difficult to classify using both statistical and pedagogical methods, for example, "relate," "prepare," and "compare." These BTAVs are ambiguous in the real sense. These verbs appear in three or more (i.e., ≥50%) of the cognitive levels and have occurrence frequency greater than the threshold value. Examples of such words are "relate" occurring in *Knowledge, Application*, and *Analysis*; "prepare" occurring in *Application, Evaluation*, and *Create*; "compare" occurring in *Comprehension, Analysis*, and *Evaluation.* Table 9 shows the list of ambiguous verbs.

Table 7. List of Single-Level Transitional Verbs.

Action verb	Knowledge (*f*)	Comprehension (*f*)
Describe	*14*	*22*
Identify	*20*	*14*
Locate	*10*	*10*
recognize	*14*	*11*
Action verb	Comprehension (*f*)	Application (*f*)
interpret	*17*	*15*
Action verb	Application (*f*)	Analysis (*f*)
discover	*8*	*4*
experiment	*6*	*9*
Action verb	Analysis (*f*)	Evaluation (*f*)
appraise	*11*	*22*
criticize	*11*	*11*

Table 8. List of Multi-Level Transitional Verbs.

Action verb	Knowledge (*f*)	Analysis (*f*)
outline	*11*	*10*
select	*16*	*12*
order	*5*	*5*
Action verb	Comprehension (*f*)	Analysis (*f*)
distinguish	*12*	*21*
infer	*15*	*14*
classify	*18*	*10*
Action verb	Comprehension (*f*)	Evaluation (*f*)
defend	*12*	*15*
estimate	*11*	*15*
Action verb	Application (*f*)	Evaluation (*f*)
choose	*11*	*10*
Action verb	Application (*f*)	Create (*f*)
construct	*13*	*29*
modify	*12*	*10*
produce	*13*	*13*

Result

Result of Level 1 (Unique BTAVs With 10 or More Occurrence Frequency)

It was observed that 25 BTAVs based on the question were found in Level 1, out of which 22 (88%) matched with the gold standard. Three BTAVs that did not match with the gold standard are show, use, and arrange. The error analysis is explained in Section "Discussion and Error Analysis." Table 10 shows the list of BTAVs that occurred in the question set in Level 1.

Result of Level 2 (Unique BTAVs With Less Than 10 Occurrence Frequency)

It was observed that eight BTAVs based on the question were found in Level 2, out of which seven (87.5%) matched with the gold standard. The BTAV that did not match with the gold standard is "*collect*" The error analysis is explained in Section "Discussion and Error Analysis." Table 11 shows the list of BTAVs that occurred in the question set in Level 2.

Table 9. List of Ambiguous Verbs.

Action verb	Knowledge (*f*)	Application (*f*)	Analysis (*f*)
Relate	*11*	*12*	*17*
Action verb	Application (*f*)	Evaluation (*f*)	Create (*f*)
prepare	*11*	*12*	*12*
Action verb	Comprehension (*f*)	Analysis (*f*)	Evaluation (*f*)
compare	*11*	*24*	*18*

Table 10. List of BTAVs That Occurred in the Question Set in Level 1.

Knowledge	Comprehension	Application	Analysis	Evaluation	Create
Cite	convert	Act	*analyze*	Argue	*arrange*
Define	*discuss*	*Apply*	*categorize*	Assess	Assemble
Label	*explain*	Calculate	contrast	*Conclude*	Combine
List	*express*	Compute	*diagram*	critique	compose
Match	extend	Demonstrate	differentiate	*evaluate*	*create*
memorize	generalize	Dramatize	discriminate	judge	*design*
Name	paraphrase	Employ	divide	manage	develop
Recall	predict	*Illustrate*	*examine*	rearrange	devise
Recite	report	Manipulate	point out	reconcile	formulate
Record	restate	Operate	question	set up	generate
Repeat	review	Practice	*separate*	synthesize	invent
reproduce	rewrite	Schedule	subdivide		organize
State	*summarize*	*Show*	test		plan
	translate	Sketch			rate
		Solve			revise
		Use			*write*

Note. Italic words refer to BTAVs in question that are correct. Bold italic words refer to BTAVs in question that are incorrect. BTAVs = Bloom's Taxonomy action verbs.

Result of Level 3 (Transitional BTAVs)

The BTAVs that were identified in Level 3 are "describe," "identify," "locate," "outline," "distinguish," "classify," "choose," and "order." All the three single-level transitional verbs "describe," "identify," and "locate" were correctly mapped either in the *Knowledge* or *Comprehension* level, except once for "describe." For multi-level transitional verbs, except one verb ("choose"), which was mapped to *Knowledge* level, the rest four "outline," "distinguish," "classify," and "order" were correct.

Result of Level 4 (Ambiguous Action Verbs)

Only one action verb *compare* mapped to the *Analysis* level was identified in Level 4, and it was found to be correct.

Discussion and Error Analysis

Overall Result

The proposed classification model was successful in the identification of cognitive level as per the gold standard in 81 out of 101 questions (80.19%) in Economics, 44 out of 52 questions (84.61%) in Geography, and 29 out of 31 questions (93.54%) in Science (Biology). The proposed unambiguous classification of BTAVs in Level 1 and Level 2 had an accuracy of 88% and 87.5%, respectively.

Common Error Patterns

Out of the 30 questions where the classification model was unsuccessful, some patterns of error were observed, which are as follows.

Multiple BTAVs present in the question stem. A typical pattern for mismatch that was observed was in cases of multiple BTAVs presents in the question. For example, question ID E1-11 having BTAVs "conclude" and "illustrate" together, question ID G6-4 having BTAVs "read," "collect," and "discuss" together.

- Question E-11: Suppose records show that the average income in a country has been increasing over a period of time. From this, can we conclude that all

Table 11. List of BTAVs That Occurred in the Question Set in Level 2.

Knowledge	Comprehension	Application	Analysis	Evaluation	Create
duplicate	associate	*Change*	breakdown	*decide*	compile
Quote	characterize	*Complete*	correlate	grade	facilitate
Read	give examples	Backup	deduce	weigh	hypothesize
Tabulate	*indicate*	Implement	dissect	counsel	integrate
Copy	represent	Interview	prioritize	mediate	originate
Draw	clarify	Paint	survey	probe	propose
underline	extrapolate	Utilize	break	release	role-play
	Give	Adapt	detect	supervise	improve
	interpolate	Simulate	diagnose	verify	make
	articulate		figure	attach	specify
	observe		inspect	core	tell / tell why
			inventory	*determine*	***collect***
			investigate	value	reconstruct
			debate		reorganize
			group		

Note. Italic words refer to BTAVs in question that are correct. Bold italic words refer to BTAVs in question that are incorrect. BTAVs = Bloom's Taxonomy action verbs.

sections of the economy have become better? Illustrate your answer with an example.
- Question G6-4: Have you read about the Kalinganagar controversy? Collect information from different sources and discuss.

Words with different cognition level with reference to the context. Another pattern that was observed was in the questions that had the BTAV "choose." The questions with the word "choose" were multiple choice questions where the learners needed to remember the correct option out of the given choices. Thus, for all such choose cases (4 times in Economics), the questions are marked unsuccessful as they were neither given *Application* or *Evaluation* level as per the classification. A similar example is the word "arrange."

- Question E4-37: Choose the most appropriate answer.
- Question E5-12: Arrange the following in the correct order.

Words related to physical activity. Words that involved a physical cognition such as "collect", "use" and "show" were found to be mismatched. The reason is that these words can be applied to any level depending upon reference to the context. So unless the context is considered, these words tend to be ambiguous.

- Question E1-22: In Tamil Nadu, 75% of the people living in rural areas use a ration shop, whereas in Jharkhand only 8% of rural people do so. Where would people be better off and why?
- Question G3-4: Collect information on how industries are polluting our water.
- Question G5-2: Superimpose the maps showing distribution of iron ore, manganese, coal and iron, and steel industry. Do you see any correlation? Why?

Words with no statistical frequency of belonging to a cognitive level. For words such as "measure," "trace," "consider," the proposed classification model does not have any statistical occurrence frequency data. The only reference is the Northwestern University data set. Thus, it is difficult to predict the cognitive level of such words.

- Question E1-27: In what respects is the criterion used by the United Nations Development Programme (UNDP) for measuring development different from the one used by the World Bank?
- Question E5-20: What factors gave birth to the consumer movement in India? Trace its evolution.
- Question B6-4: What processes would you consider essential for maintaining life?

Words belonging to transitional verbs. Words belonging to transitional verb once resulted in a cognitive level mismatch (i.e., mapped to another different cognitive level than the transitional levels). For example, the word "Describe" can be in both *Knowledge* and *Comprehension* levels. But once it was mapped to *Analysis* level.

- Question G4-13: Describe the impact of globalization on Indian agriculture.

Implications of the Findings

The proposed research work contributes the following—First, a compilation of 153 unique Bloom's Taxonomy

cognitive level action verbs, which is based on statistical ranking and validated on data sets of three different school textbook questions. Also, this is verified based on the gold standard provided by human annotation. This compilation is free from overlapping in multiple cognitive levels (making it less ambiguous) and helps users identify the necessary cognitive level associated with the word. Second, a separate compilation of 21 action verbs is provided, which shows a transition from one cognitive level to another. Thus, the ambiguity of the transitional verbs is reduced to the specific cognitive levels. Third, a few ambiguous action verbs were identified that should be avoided.

The proposed research reduced the cognitive level's ambiguity in the action verbs and made Bloom's Taxonomy classification more robust. This makes it easier for teachers to create and set the cognitive levels previously affected by BTAVs' ambiguity.

Conclusion

Krathwohl (2002) argues that nearly all complex learning activities require the use of several different cognitive skills. Like any theoretical model, "Bloom's Taxonomy has its strengths and weaknesses. Its greatest strength is that it has taken the very important topic of thinking and placed a structure around it that is usable by practitioners." There is no doubt that those teachers who keep a list of questions set related to the different levels of Bloom's Taxonomy do a better job of encouraging higher-order thinking in their students than those who do not have such a tool. On the contrary, as someone who has collaborated with a group of educators to identify a set of taxonomy-specific questions and learning activities can testify, "there is little agreement about what self-evident words such as 'study' or 'evaluation' mean." Furthermore, too many worthwhile tasks, such as real problems and ventures, cannot be mapped to the taxonomy, and attempting to do so will minimize their potential as opportunities for learning.

In the taxonomy, the level of knowledge consists of both noun and verb forms. While the target component described as a noun form is located in the full-frame bottom steps of the knowledge stage, verb forms describing the cognition process are identified as identifying and remembering the information by the students. Consequently, the stage of knowledge which is supposed to have two-dimensional characteristics becomes unilateral. Within the framework of the cognition method, the unilateral structure of taxonomy fails.

The revised taxonomy also states that Bloom's Taxonomy's cognitive domain has an additive sort of structure. It steps forward based on the degree of difficulty, and according to the need, to activate a former one for the next step. The taxonomy presents its cognitive process in categories. They are different from each other, just in terms of difficulty. However, the rigid hierarchy between categories was later softened, and category overlaps were provided (Krathwohl, 2002). For this reason, combining program objectives, teaching, and assessment is more crucial than ever, as mentioned by Pickard (2007). Bloom's revised Taxonomy did not introduce a drastic improvement to Bloom's original classification, though it has provided some crucial developments.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Syaamantak Das https://orcid.org/0000-0001-9896-3312

Notes

1. The article by Stanny states it has 176 action verbs. The actual number of action verbs is 177. The word tell is merged with tell/tell why. Erratum acknowledged by the author. For this research, it is considered two different words.
2. http://www.northeastern.edu/nuolirc/wp-content/uploads/2018/01/Blooms-Taxonomy-Handout.pdf

References

Abdelrahman, M. (2014). An analysis of the tenth grade English language textbooks questions in Jordan based on the revised edition of Bloom's Taxonomy. *Journal of Education and Practice*, *5*(18), 139–151.

Agrawal, R., Gollapudi, S., Kannan, A., & Kenthapadi, K. (2012). Electronic textbooks and data mining. In H. Gao, L. Lim, W. Wang, C. Li, & L. Chen (Eds.), *Web-age information management*. Lecture Notes in Computer Science, Vol. 7418. Springer, Berlin. https://doi.org/10.1007/978-3-642-32281-5_3

Agrawal, R., Gollapudi, S., Kannan, A., & Kenthapadi, K. (2014). Study navigator: An algorithmically generated aid for learning from electronic textbooks. *Journal of Educational Data Mining*, *6*(1), 53–75.

Allan, J. (1996). Learning outcomes in higher education. *Studies in Higher Education*, *21*(1), 93–108.

Assaly, I. R., & Smadi, O. M. (2015). Using Bloom's Taxonomy to evaluate the cognitive levels of master class textbook's questions. *English Language Teaching*, *8*(5), 100–110.

Bayerl, P. S., & Paul, K. I. (2011). What determines inter-coder agreement in manual annotations? A meta-analytic investigation. *Computational Linguistics*, *37*(4), 699–725.

Bloom, B. S. (1956). *Taxonomy of educational objectives: Vol. 1. Cognitive domain* (pp. 20–24). McKay.

Dalton, J., & Smith, D. (1989). *Extending children's special abilities: Strategies for primary classrooms*. Office of Schools Administration, Ministry of Education.

DeWaelsche, S. A. (2015). Critical thinking, questioning and student engagement in Korean university English courses. *Linguistics and Education*, *32*, 131–147.

Fink, L. D. (2013). *Creating significant learning experiences: An integrated approach to designing college courses*. John Wiley.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34.

Igbaria, A. K. (2013). A content analysis of the wh-questions in the EFL textbook of "horizons." *International Education Studies*, *6*(7), 200–224.

Jones, K. O., Harland, J., Reid, J. M. V., & Bartlett, R. (2009). Relationship between examination questions and bloom's taxonomy. In *39th IEEE Frontiers in Education Conference*, pp. 1–6. San Antonio. https://doi.org/10.1109/FIE.2009.5350598.

Krathwohl, D. R. (2002). A revision of Bloom's Taxonomy: An overview. *Theory into Practice*, *41*(4), 212–218.

Krathwohl, D. R., & Anderson, L. W. (2010). Merlin C. Wittrock and the revision of Bloom's Taxonomy. *Educational Psychologist*, *45*(1), 64–65.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Lee, Y.-J., Kim, M., & Yoon, H.-G. (2015). The intellectual demands of the intended primary science curriculum in Korea and Singapore: An analysis based on revised bloom's taxonomy. *International Journal of Science Education*, *37*(13), 2193–2213.

Lee, Y.-J., Kim, M., Jin, Q., Yoon, H.-G., & Matsubara, K. (2017). Revised Bloom's Taxonomy—The Swiss army knife in curriculum research. In *East-Asian primary science curricula*. Springer Briefs in Education. Springer, Singapore. https://doi.org/10.1007/978-981-10-2690-4_2

Otter, S. (1995). Learning outcomes in higher education. In J. Burke (Ed.), *Outcomes, learning and the curriculum: Implications for NVQs, GNVQs and other qualifications* (p. 273). Falmer Press.

Pickard, M. J. (2007). The new Bloom's Taxonomy: An overview for family and consumer sciences. *Journal of Family and Consumer Sciences Education*, *25*(1), 45–55.

Plack, M. M., Driscoll, M., Marquez, M., Cuppernull, L., Maring, J., & Greenberg, L. (2007). Assessing reflective writing on a pediatric clerkship by using a modified Bloom's Taxonomy. *Ambulatory Pediatrics*, *7*(4), 285–291.

Riazi, A. M., & Mosalanejad, N. (2010). Evaluation of learning objectives in Iranian high-school and pre-university english textbooks using Bloom's Taxonomy. *TESL-EJ*, *13*(4), 1–16.

Sheldon, L. E. (1988). Evaluating ELT textbooks and materials. *ELT Journal*, *42*(4), 237–246.

Stanny, C. (2016). Reevaluating Bloom's Taxonomy: What measurable verbs can and cannot say about student learning. *Education Sciences*, *6*(4), Article 37.

Stelzer, T., Gladding, G., Mestre, J. P., & Brookes, D. T. (2009). Comparing the efficacy of multimedia modules with traditional textbooks for learning introductory physics content. *American Journal of Physics*, *77*(2), 184–190.

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, *9*(4), 1–19.

Swart, A. J., & Daneti, M. (2019). Analyzing learning outcomes for electronic fundamentals using Bloom's Taxonomy. In *IEEE Global Engineering Education Conference (EDUCON)*, pp. 39–44. Dubai, United Arab Emirates. https://doi.org/10.1109/EDUCON.2019.8725137.