Short Communication

Assessing the inter-rater reliability and accuracy of pharmacy faculty's Bloom's Taxonomy classifications

Samuel C. Karpen, PhD[*], Adam C. Welch, PharmD

Bill Gatton College of Pharmacy, East Tennessee State University, Johnson City, TN

ARTICLE INFO

ABSTRACT

Objective: To identify inter-rater reliability and accuracy of pharmacy faculty members' classification of exam questions based on Bloom's Taxonomy.

Methods: Faculty at a college of pharmacy was given six example exam questions to assign to the appropriate Bloom's level.

Results: Inter-rater reliability and accuracy were both low at 0.25 and 46.0%, respectively. Accuracy increased to 81.8% when the six Bloom's levels collapsed to three.

Conclusions: Both inter-rater reliability and accuracy were low. Faculty members' misclassifications suggested a three-tier combination of the Bloom's levels that would optimally improve accuracy: Knowledge, Comprehension/Application, and Analysis/Synthesis/Evaluation. Faculty development should also be considered in improving accuracy and reliability.

Introduction

The Accreditation Council for Pharmacy Education (ACPE) recently revised its accreditation standards to focus on both, students' development of professional knowledge, skills, abilities, behaviors, and attitudes, and the manner in which student development is assessed.[1] A tool that faculty will likely use to assess student development is the cognitive component of Bloom's Taxonomy. Indeed, a ProQuest search for "Bloom's Taxonomy" returned 4160 publications and a Google search yielded 513,000 hits, an impressive testament to its widespread usage. Furthermore, pharmacy education researchers have encouraged its usage to assess student's critical thinking abilities.[2,3]

Published in 1956, Bloom's Taxonomy describes a hierarchy of six categories in the cognitive domain (knowledge, comprehension, application, analysis, synthesis, and evaluation) that progress from basic to advanced levels of learning, such that advancement to higher levels depends on mastery of lower levels.[4] Students cannot apply information that they do not comprehend, or comprehend information of which they are unaware. See Table 1 for a complete description of each level of Bloom's Taxonomy. As such, Bloom's Taxonomy provides a valuable framework for classifying the breadth of learning objectives across a variety of categories in a cognitive domain. While the taxonomy has maintained its hierarchical structure, it has undergone several revisions and extensions, including application in the affective and psychomotor domains, and a 2002 revision that better aligned the levels with their intended outcomes. The revised levels are remember, understand, apply, analyze, evaluate, and create.[5]

Regardless of its application or form, Bloom's Taxonomy is only informative if educators use it accurately and reliably. Indeed, pharmacy faculty often map course content to specific programmatic outcomes using Bloom's Taxonomy.[6,7] If faculty members interpret Bloom's Taxonomy in different ways, the reliability of content maps may be compromised.

Plack et al.[8] addressed faculty's ability to reliably use a modified version of Bloom's Taxonomy with three levels—Knowledge/

* Corresponding author: Samuel C. Karpen, PhD, Bill Gatton College of Pharmacy, East Tennessee State University, PO Box 70413, Johnson City, TN 37614.
E-mail address: karpen@etsu.edu (S.C. Karpen).

Table 1
Descriptions of the original six levels of Bloom's Taxonomy are provided below

Category	Competency
Knowledge	Define terms or recall course material
Comprehension	Explain or communicate course material in their own words
Application	Apply course material in concrete situations
Analysis	Deconstruct a topic, idea, or process into its component parts
Synthesis	Combine the parts of a topic, idea, or process in new and useful ways
Evaluation	Judge the value or legitimacy of positions, ideas, or approaches

Comprehension, Application/Analysis, and Synthesis/Evaluation and found that inter-rater reliability ranged from κ = 0.57 to 0.73 depending on the pair of raters being compared. A Kappa statistic of about 0.60 is generally considered acceptable, it appears that faculty can use a modified version of Bloom's Taxonomy rather reliably.[9] While Plack's results are promising, one must ask whether faculty would have performed as well using all six levels, and whether faculty's classifications were accurate. Consequently, we ask (1) whether college of pharmacy faculty can reliably classify questions to the original six levels of Bloom's Taxonomy, as indicated by inter-rater reliability; (2) whether pharmacy faculty can accurately classify questions to the original six levels of Bloom's Taxonomy, as indicated by percent of questions classified correctly; and (3) how do faculty most commonly misclassify questions?

Methods

An online pilot survey that required participants to categorize six exam questions—one question for each of six levels of Bloom's Taxonomy—was developed in Formstack® (Formstack®, Indianapolis, IN) and administered to all faculty at Bill Gatton College of Pharmacy. The survey was administered to 27 faculty members (18 in pharmacy practice and nine in pharmaceutical sciences). The exam questions, which were written to be clear examples of each level of Bloom's Taxonomy, were taken from the Teacher Resources section of the University of California at Berkley's Center of Teaching and Learning website.[10] See Table 2 for a list of questions and their corresponding level. After assigning each question to a Bloom's Taxonomy level, faculty rated their overall confidence in their classification on a 5-point Likert scale (1 = not confident at all and 5 = extremely confident). Krippendorff's alpha, for which 0.60 is generally considered acceptable reliability, was used to determine the inter-rater reliability of faculty's classifications.[11,12] Additionally, we determined the accuracy of their classifications (percent correct) and most common misclassifications. The most common misclassifications were used to determine the optimal collapsing of categories, should colleges of pharmacy wish to combine some of the often confused categories to reduce burden on faculty who are required to classify their exam items to Bloom's Taxonomy. This study was exempt from review as determined by the investigators' Institutional Review Board.

Results

There were 21 (77.8%) faculty members who participated in the survey (14 in pharmacy practice and 7 in pharmaceutical sciences).

Inter-rater reliability

The combined inter-rater reliability for all faculty members was α = 0.25 (95% CI: 0.21–0.29) for the six questions.

Accuracy

Overall accuracy was 46.0%. Raters did not misclassify questions randomly; rather, misclassifications usually fell in nearby categories. Raters classified Knowledge questions quite accurately, but often confused Comprehension questions with Application questions, and Analysis, Synthesis, and Evaluation questions with one another. Table 3 describes the classifications that the participants made for each question. This pattern suggests that if one were to modify Bloom's Taxonomy for question tagging

Table 2
The questions used in the study were posted in the University of California at Berkley's School of Education website and were written to be examples of the six Bloom's levels

Category	Question
Knowledge	Write the equation for the ideal gas law
Comprehension	Translate the following paragraph from "Der Spiegel" into good English
Application	Determine the centroid of a plane figure
Analysis	Given an argument for the abolition of guns, enumerate the positive and negative points presented
Synthesis	Given two opposing theories design an experiment to compare them
Evaluation	In a given clinical situation, select the most reasonable intervention and predict the main effects and possible side effects