CENTRE DE RECHERCHE SUR LES STRATÉGIES ÉCONOMIQUES

Examination design: an axiomatic approach

Sylvain Béal, Sylvain Ferrières

September 2019

Working paper No. 2019 – 05

CRESE
30, avenue de l'Observatoire
25009 Besançon
France
http://crese.univ-fcomte.fr/

UNIVERSITÉ DE FRANCHE-COMTÉ

Examination design: an axiomatic approach[☆]

Sylvain Béal[a,*], Sylvain Ferrières[b]

[a] *CRESE EA3190, Univ. Bourgogne Franche-Comté, F-25000 Besançon, France*
[b] *Université de Saint-Etienne, CNRS UMR 5824 GATE Lyon Saint-Etienne, France*

Abstract

We use the axiomatic approach for the design of examination. Each course is identified with a set of learning objectives, and an exam is composed of exercises each of which is identified with the learning objectives that it assesses. Our results characterizes several well-known structures for the exams in terms of the assignment of learning objectives: partitions, symmetric balanced collections, orderings.

Keywords: Examination, learning objective, axiomatic approach, partitions, balanced collections, orderings, consistency.

1. Introduction

For decades, there has been a consensus on the need classroom evaluation, and this assessment is intrinsically linked to student learning and performance. Assessment often goes hand in hand with grading. Walvoord and Johnson Anderson (2009) refers to grading as "the process by which a teacher assesses student learning through classroom tests and assignments". In this article, we would like to study the design of an evaluation/test from an axiomatic point of view.

The field of educological research includes many studies on the links between teaching, learning and grading. Walvoord (2010) argues that the first step of assessment is to determine what the student should be able to do after completing a course. These goals may also be called learning outcomes or learning objectives and are incorporated to the course description in many universities nowadays. In this framework, two areas of research are particularly fertile. Firstly, the literature sometimes called "curriculum mapping" makes it clear how does assessment relate to teaching program. According to Harden (2001),

[☆] We would like to thank Francis Bloch and the participants at the 2018 Paris symposium on games for useful comments. Financial support from research programs IDEXLYON from Université de Lyon (project INDEPTH) within the Programme Investissements d'Avenir (ANR-16-IDEX-0005), and "Mathématiques de la décision pour l'ingénierie physique et sociale" (MODMAD) is gratefully acknowledged.

[*] Corresponding author

Email addresses: `sylvain.beal@univ-fcomte.fr` (Sylvain Béal), `sylvain.ferrieres@univ-fcomte.fr` (Sylvain Ferrières)

URL: `https://sites.google.com/site/bealpage/` (Sylvain Béal), `https://sites.google.com/site/sylvainferrieres/` (Sylvain Ferrières)

"Curriculum mapping is concerned with (...) the measures used to determine whether the student has achieved the expected learning outcomes (assessment)."

He states that in medical education, the map identifies the learning outcomes assessed in each of the questions in a multiple choice question test. Secondly, in educational psychology, Bloom's taxonomy (Bloom, 1956) is a famous set of hierarchical models used to classify educational learning objectives into levels of complexity and specificity. Airasian and Miranda (2002) point out that this taxonomy has been widely used to provide "*more clearly defined assessments and a stronger connection of assessment to both objectives and instruction*". Here, "instruction" refers to the material taught during a course and "objectives" refers to the teaching objectives/learning objectives. Airasian and Miranda (2002) insist on the fact that

"(...) the information obtained during the assessment process is influenced to a great extent by what has preceded it during the instructional process, particularly as both processes (instruction and assessment) are aligned with the stated objective. If the three components are well aligned, the assessment results are likely to be reasonably valid."

In this article, we introduce a simple model of examination design which tries to incorporate these ideas. More specifically, a course is identified with a set of learning objectives (indifferently, it can be chapters, concepts, etc.). For instance, a classical course in noncooperative game theory can include the following learning objectives: Normal form games, Extensive form games, Dominant strategies, Nash equilibrium (NE), Mixed strategy NE, Information sets, Backward induction, Subgame perfect NE, Repeated games, Folk Theorems, Incomplete information games, Bayesian NE. A written exam is then represented by a set of exercises, each of which can be identified with the learning objectives that the students must master to provide the right answer. As an example, if an exercise requires to compute the set of mixed Nash equilibria after removing iteratively all strictly dominated strategies in a given two-player matrix game, then it can be identified with the set of learning objectives $\{\text{Dominant strategies}, \text{Mixed Nash NE}\}$.

Our main objective is to study the mapping between a course (a set of learning objectives) and an exam (a set of exercises assessing various learning objectives). We call such a mapping an examination function. Therefore, we focus on the interplay between learning objectives and exams by considering that a course can be summarized by its associated learning objectives. In order to delineate the shape of examination functions, we introduce several axioms (or properties) that we believe are reasonable for the design of a relevant examination. These axioms can help to answer questions such as: should an ideal exam cover all the content of a course? How can an exam determine whether a learning objective has been achieved? Should each learning objective be evaluated the same number of times? A first axiom, called Full covering, requires that any learning objective is assessed by at least one exercise. Another axiom, called Non redundancy requires that each exercise assesses at least one learning objective exclusively. We also invoke two natural axioms of consistency, which discuss the changes in the exam structure if a learning objective is removed (because the corresponding part in the course could not be taught). In such a situation, the axiom of Sharp consistency imposes to keep only the exercises that do not evaluate the deleted learning objectives, while the axiom of Accommodating consistency requires adapting existing exercises to not evaluate the deleted learning objective.

Our results identify several well-known structures. Proposition 1 characterizes the examination functions in which the learning objectives are partitioned into exercises by means of Full covering, Non redundancy, Accommodating consistency and the axiom of Balancing. The later axiom imposes that each learning objective is assessed in the same number of exercises. Propositions 2 and 3 single out the finest partition, that is, the exam containing one exercise for assessing exclusively each learning objective. Proposition 4 characterizes exams in which the structure of learning objectives is a symmetric balanced collection. Balanced collections are used in cooperative game theory for proving the non-emptiness of the core (Bondareva, 1963; Shapley, 1967). In Proposition 5, there is an ordering over the learning objectives, which reflects prerequisites. In an exam, the learning objectives are gradually added so as to constitute exercises with progressive difficulty, provided that the difficulty of an exercise can be measured by the number of learning objectives it assesses.

From a mathematical point of view, our approach consists in assigning to each set of objects a nonempty set of subsets of these objects. This approach is sometimes known as block design in combinatorial mathematics, where the learning objectives are called points and the exercises are called blocks (see Wallis, 2007, for instance). If the objects are players, then subsets of objects are coalitions of players and so an examination function would assign to each set of players a collection of coalitions of players. This analogy is present in van den Brink (2012), which can be considered as the closest article to our work. This article characterizes the sets of feasible coalitions in communication networks and compares them with sets of feasible coalitions arising from hierarchies (see also Bilbao, 2000, for other combinatorial structures which can be associated with cooperative games). Beyond a completely different interpretation, our approach departs from van den Brink (2012) in two directions. Firstly, our set of objects (learning objectives) can vary, and some key axioms of consistency actually describe what happens in such a case. To the contrary, the set of objects (the player set) in van den Brink (2012) is fixed. Secondly, the collections of subsets of objects that are characterized in the two articles as well as most of the invoked axioms are very different too. Let us mention that the axiomatic approach is also used in the field of education by Zapechelnyuk (2015), who characterizes a specific scoring rule for multiple-choice tests.

We are aware of the limits of our approach. We do not consider a priori relations among the set of learning objectives. It would be more realistic to assume a dependence relation specifying that mastering some given learning objectives is a prerequisite to master another learning objective. For instance, in our example above, it makes sense to assume that Extensive form games, Information sets and Backward induction are prerequisite to master Subgame perfect NE. The unique exception is Proposition 5 in which an implicit ordering over the learning objectives emerges endogenously. We neither consider that different learning objectives may be of different importance. We also leave aside the issue of exam grading. The last section of the article briefly takes a look at some of these aspects.

The rest of the article is organized as follows. Section 2 introduces the model. Section 3 presents the axioms that we impose on an examination function. Section 4 states the results. Their proofs and the demonstration of the logical independence of the axioms are relegated to the Appendix. Section 5 provides concluding remarks.

2. A simple model of examination design

Let $U \subseteq \mathbb{N}$ be a finite universe of at least 3 **learning objectives** that a teacher want to assess. These learning objectives can be represented by chapters, definitions, concepts, theorems or any other component that may be presented during a course. In other words, we summarize a course by a set of learning objectives. Any nonempty subset N of U is thus called a **course**. We denote by 2^N the power set of N, that is the set of all subsets of learning objectives taught during course N. Also, we write u and n for the cardinality of sets $|U|$ and $|N|$, and so on.

In order to assess the learning objectives associated with course N, a teacher has to design a final exam. This written exam takes the form of a set of exercises, and we identify each exercise by the set of learning objectives it assesses. A **final exam** is thus a collection of subsets of learning objectives, i.e. for a course N, it is an element of 2^{2^N}. We are interested in a function which systematically assigns a final exam to each conceivable course, and we rule out final exams containing an empty exercise, i.e. we restrict ourselves to elements of $2^{2^U \setminus \varnothing}$. Formally, we look for a single-valued **examination function** function $f : 2^U \longrightarrow 2^{2^U \setminus \varnothing}$, which assigns to each nonempty course $N \in 2^U$ a final exam $f(N) \in 2^{2^N \setminus \varnothing}$, i.e. a collection of nonempty subsets of the set of learning objectives N. For an examination function f, a course N and any learning objective $i \in N$, we say that i is **covered** by f at N if there is $S \in f(N)$ such that $S \ni i$.

Some examples of examination functions are the following.

1. The function f such that, for each $N \in 2^U$, $f(N) = \{\{i\}, i \in N\}$ is the examination function which assigns to each course a final exam composed of n exercises, one assessing each learning objective. Each learning objective has an exclusive evaluation.
2. The function f such that, for each $N \in 2^U$, $f(N) = 2^N$ is the examination function which assigns to each course a final exam composed of 2^n exercises, one assessing each combination of learning objectives. The teacher wants to test all combinations of learning objectives, perhaps to account for potential synergies between learning objectives.
3. The function f such that, for each $N \in 2^U$, $f(N) = \{\{1, \dots, i\}, i \in N\}$ is the examination function which assigns to each course a final exam composed of n exercises, one assessing learning objective 1, one assessing learning objectives 1 and 2, one assessing learning objectives 1, 2 and 3, and so on. Here, learning objectives may be ordered from the first taught (learning objective 1) to the last taught (learning objective n). The difficulty increases gradually by assessing a new learning objective with each new exercise.
4. The function f such that, for each $N \in 2^U$, $f(N) = \{N\}$ is the examination function which assigns to each course a final exam composed of a single exercise assessing simultaneously all learning objectives. The teacher only wants to assess learning objectives altogether and not separately.

Remark that the function f such that, for each $N \in 2^U$, $f(N) = \varnothing$ is allowed, whereas the function f such that, for each $N \in 2^U$, $f(N) = \{\varnothing\}$ is ruled out according to the definition of an examination function. In other words, we allow for empty exams, but not for nonempty exams containing an empty exercise.

For any integer $k \in \{1, \ldots, u\}$ and any course $N \in 2^U$, we denote by $C(k, N)$ the set of all subsets of N of size k, i.e.

$$C(k, N) = \{S \in 2^N : s = k\}.$$

Beyond the few above examples, there are other well-known types of structure for an examination function. A **partition** of N is a collection P of subsets of N such that any two of its elements are disjoint and their union is N. In particular, our aforementioned examples 1 and 4 induce the finest and coarsest partitions of N, and can be formulated as $C(1, N) = \{\{i\}, i \in N\}$ and $C(n, N) = \{N\}$, respectively. A **balanced collection** (Shapley, 1967) for N is a collection B of subsets of N such that there is a non-negative weight λ_S for each $S \in B$ and, for each $i \in N$,

$$\sum_{S \in B : S \ni i} \lambda_S = 1.$$

Partitions are special instances of balanced collections with weights all equal to one.

3. The axioms

Below, we list some axioms for an examination function f. The first axiom imposes that each learning objective of a given course is assessed in at least one exercise or, equivalently, that all learning objectives are covered by f.

Full covering (FC). For each nonempty $N \in 2^U$, $\bigcup_{S \in f(N)} S = N$.

Full covering is equivalent to the axiom of Normality in van den Brink (2012). The second axiom is weaker: it requires that an exam always contains at least one exercise.

Nonemptiness (N). For each nonempty $N \in 2^U$, $f(N) \neq \varnothing$.

Another axiom deals with the minimality of the final exam in the sense that removing an exercise should end up in a situation in which some learning objective is not covered. In other words, every exercise assesses at least one learning objective exclusively, otherwise it should be deleted. Observe that this requirement does not imply that all learning objectives are covered in the original situation.

Non redundancy (NR). For each nonempty $N \in 2^U$ and each $S \in f(N)$, $\bigcup_{T \in f(N) \setminus \{S\}} T \neq N$.

As an example, consider the course $N = \{1, 2, 3, 4\}$. The exam $f(N) = \{\{1, 2, 3\}, \{1, 3, 4\}\}$ satisfies Non redundancy on N: exercise $\{1, 2, 3\}$ assesses exclusively learning objective 2 (meaning that 2 would not be covered by f at N without exercise $\{1,2,3\}$) and exercise $\{1,3,4\}$ assesses exclusively learning objective 4. The exam $f(N) = \{\{1,2\},\{1,3\},\{2,3,4\}\}$ violates Non redundancy on N: without exercise $\{1,2\}$, the final would still cover all learning objectives.

We now introduce two variants of the popular principle of consistency (see Thomson, 2019). In both cases, we discuss the consequences of deleting a learning objective on the structure of the final exam. In other word, if the course has less content, how should the final exam be adapted? In the first axiom, we impose that the adjusted final exam consists of all the originally selected exercises that do not assess the deleted learning objective. The rationale behind this principle is that each exercise is seen as an indivisible entity, which means that the exercise has no meaning if one of the associated learning objectives is not assessed anymore. In the second axiom, to the contrary, we require that the adjusted final exam contains not only the originally selected exercises that do not include the deleted learning objective but also the exercises such that the addition of the deleted learning objective yields an originally selected exercise. These two visions are somehow similar to the γ and β approaches examined by Hart and Kurz (1983) for the stability of coalition structures.

Sharp consistency (SC). For each nonempty $N \in 2^U$ and each $i \in N$, $f(N\backslash\{i\}) = \{S \in f(N) : S \not\ni i\}$.

Accommodating consistency (AC). For each nonempty $N \in 2^U$ and each $i \in N$, $f(N\backslash\{i\}) = \{(S\backslash\{i\}) \neq \varnothing : S \in f(N)\}$.

For a given course $N = \{1,2,3,4\}$, suppose that $f(N) = \{\{1\},\{1,4\},\{2,3,4\},\{1,3\},\{3\}\}$. Then, on $N\backslash\{1\}$, f satisfies the requirement of Sharp consistency if $f(N\backslash\{1\}) = \{\{2,3,4\},\{3\}\}$ and the requirement of Associated consistency if $f(N\backslash\{1\}) = \{\{4\},\{2,3,4\},\{3\}\}$. In the latter exam, exercise $\{3\}$ comes from $f(N)$ by applying Associated consistency to exercise $\{1,3\}$ and also because $\{3\}$ belongs to $f(N)$.

Next, we introduce two axioms which compare the treatment of two learning objectives. The first one imposes that each learning objective is assessed in the same number of exercises. The second one imposes a close principle by relying on endogenous weights.

Balancing (B). For each nonempty $N \in 2^U$ and each $i,j \in N$, $|\{S \in f(N) : S \ni i\}| = |\{S \in f(N) : S \ni j\}|$.

Weighting (W). For each nonempty $N \in 2^U$ and each $i,j \in N$, $\sum_{S \in f(N): S \ni i} 1/s = \sum_{S \in f(N): S \ni j} 1/s$.

In the Weighting axiom, the importance of each exercise is in a sense inversely proportional to the number of learning objectives it assesses, while it is not the case in the Balancing axiom. The two axioms are not logically related to each other. For any $i \in U$, an examination function f such that $f(N) = \{\{i\}, N\backslash\{i\}\}$ satisfies **B** but not **W**. For any triple of elements $\{i,j,k\} \subseteq N$, an examination function f such that $f(N) = C(2,\{i,j,k\}) \cup C(1, N\backslash\{i,j,k\})$ satisfies **W** but not **B**. Remark also that both axioms of Balancing and Weighting do not impose any kind of symmetry regarding the size of the exercises within the chosen exam. If $N = \{1,2,3,4\}$, then $f(N) = \{\{1\},\{1,2,3\},\{2,4\},\{3,4\}\}$ satisfies Balancing on N and $f(N) = \{\{1\},\{2,3\},\{2,4\},\{3,4\}\}$ satisfies Weighting on N.

Next, we present an axiom which aims at discriminating all learning objectives. More specifically, it imposes that for each learning objective, there is not another learning objective assessed in all exercises evaluating the first learning objective.

Discrimination (D). For each nonempty $N \in 2^U$ and each $i \in N$, $\bigcap_{S \in f(N) : S \ni i} S \subseteq \{i\}$.

A plausible interpretation is that for any learning objective i, there is no other learning objective j that can interfere with the assessment of i. To see this, consider the course $N = \{1,2,3,4\}$. The exam $f(N) = \{\{1\},\{1,2\},\{1,2,4\},\{1,3\},\{3,4\}\}$ violates Discrimination on N: learning objective 1 is systematically assessed in all exercises evaluating learning objective 2. Thus, if a teacher want to determine whether learning objective 2 has been achieved, it is not possible to rely solely on the exercises evaluating this learning objective. To the contrary, the exam $f(N) = \{\{1\},\{1,2\},\{1,2,3\},\{2,3\},\{3,4\}\}$ satisfies Discrimination on N. The next axiom attempts to capture the idea of a progressive difficulty in the exercises of an examination, if, once gain, the number of assessed learning objectives is a good proxy for the difficulty of an exercise.

Gradual complexity (GC). For each nonempty $N \in 2^U$ and each $k \in \{1,\dots,n\}$, there is $S \in f(N)$ such that $s = k$.

Gradual complexity means that there are at least n exercises in the exam, and at least one assessing k learning objectives for each size/complexity $k \in \{1,\dots,n\}$. Note that Gradual complexity does not impose that the exercises have a nested structure, even in the situation where there are just n exercises, as shown by the exam $f(N) = \{\{1\},\{2,3\},\{1,2,4\},\{1,2,3,4\}\}$ associated with a course $N = \{1,2,3,4\}$. For any examination function f satisfying Gradual complexity, any examination function f' such that, for each $N \in 2^U$, $f(N) \subseteq f'(N)$ also satisfies Gradual complexity. One way to discriminate among those examination functions is to impose the requirement below.

For a given set A of axioms, an examination function f is called **minimal with respect to axioms in** A if there is no other examination function f' satisfying the set of axioms A and such that, for each $N \in 2^U$, $|f'(N)| \leq |f(N)|$, with at least one strict inequality. This condition is desirable to avoid superfluous exercises.

4. Results

This section presents the results, which are gathered in three categories according to the structure of the selected final exam in terms of learning objectives: partitions, symmetric balanced collections and orderings.

4.1. Partitions

Our first result characterizes the examination functions such that the exam partitions the learning objectives into exercises. Moreover, partitioning is done consistently in the sense that if a learning objective is dropped, then the exam is adjusted by dropping the unique exercise assessing it.

Proposition 1. *An examination function f satisfies Full covering (**FC**), Non redundancy (**NR**), Accommodating consistency (**AC**) and Balancing (**B**) if and only if there is a partition P of U such that, for each $N \in 2^U$, $f(N)$ is the restriction of P to N, i.e.,*

$$f(N) = \{(S \cap N) \neq \varnothing : S \in P\}.$$

In order to single out the finest partition, the axiom of Discrimination can be invoked. In this case, it is even not necessary to impose Accommodating consistency and Balancing.

Proposition 2. *An examination function f satisfies Full covering (**FC**), Non redundancy (**NR**) and Discrimination (**D**) if and only if f assigns to each course N its finest partition, i.e.,*

$$f(N) = \{\{i\}, i \in N\}.$$

The next result provides another characterization of the finest partition with a completely different set of axioms. It also highlights the logical relationships between some axioms.

Proposition 3. *Consider any examination function f satisfies Nonemptiness (**N**) and Sharp consistency (**SC**). Then*

(i) for each $N \subseteq U$, $f(N) \supseteq \{\{i\}, i \in N\}$;

*(ii) f satisfies Full covering (**FC**);*

*(iii) f satisfies Non redundancy (**NR**) if and only if for each $N \subseteq U$, $f(N) = \{\{i\}, i \in N\}$;*

*(iv) f satisfies Discrimination (**D**).*

4.2. Symmetric balanced collections

A balanced collection B on a course N is called **symmetric** if for each exercise $S \in B$, all other exercises of size s are also in B. The result below is a characterization of the examination functions such that the structure of the proposed exam is a symmetric balanced collection.

Proposition 4. *An examination function f satisfies Nonemptiness (**N**), Sharp consistency (**SC**) and either Balancing (**B**) and Weighting (**W**) if and only f (consistently) assigns to each course a symmetric balanced collection containing its finest partition, i.e., there is a set $Q \subseteq \{1, \dots, u\}$ with $Q \ni 1$ such that, for each $N \in 2^U$,*

$$f(N) = \{C(q, N), q \in Q\}.$$

The reader may wonder whether it is possible to strengthen Proposition 4 to account for all balanced collections and not just the symmetric ones. This remains an open question. Studying balanced collections that are not symmetric yields serious complications given the richness of the set of balanced collections and the few identified properties of this set. An exception is the algorithm provided by Peleg (1965), which enumerates all balanced collections of order n from the set of balanced collections of order $n-1$.

4.3. Prerequisite ordering

An ordering π on U is a function $\pi : U \longrightarrow \{1, \dots, u\}$ which specifies a different position or rank $\pi(i) \in \{1, \dots, u\}$ for each learning objective $i \in U$. Denote by $P_{i,\pi} = \{j \in U : \pi(j) \le \pi(i)\}$ the set containing learning objective i as well as the learning objectives appearing before i according to π (or with a lower rank).

Proposition 5. *An examination function f satisfies Gradual complexity (**GC**), Accommodating consistency (**AC**) and is minimal with respect to these axioms if and only if there is a (prerequisite) ordering π on U such that, for each $N \in 2^U$, $f(N) = \{P_{i,\pi} \cap N, i \in N\}$.*

The examination functions characterized in the above result have a clear meaning. The ordering π on the set of learning objectives represents the timeline of the course U: the first content taught during the course is associated with learning objective $\pi^{-1}(1)$, the second content taught is associated with learning objective $\pi^{-1}(2)$, and so on. On a restricted course N, the timeline is consistently adjusted: if learning objectives i and j belongs to N and if $\pi(i) < \pi(j)$, then it is still the case when learning objectives in $U \backslash N$ (some possibly in between $\pi(i)$ and $\pi(j)$) could not be taught. As such, the ordering π on U and its restriction on N mean that i can be considered as a prerequisite to j if $\pi(i) < \pi(j)$. The examination functions studied in Proposition 5 reflect these prerequisites in the sense that there is one exercise “devoted” to each learning objective in which all its prerequisites are also assessed.

5. Concluding remarks

There are at least two natural ways to extend our study. Firstly, it would make sense to endow each learning objective with a weight reflecting its importance within a course. If $\omega_i \in \mathbb{R}_{++}$ denotes the weight associated with learning objective $i \in U$, then it is possible to replace the Full covering axiom by a weighted version relying on a kind of knowledge threshold q. More specifically, the axiom of Threshold covering can impose, for each course $N \in 2^U$, that

$$\sum_{i \in \bigcup\limits_{S \in f(N)} S} \omega_i \geq q.$$

Secondly, it would also make sense to include an exogenous prerequisite structure on the set of learning objectives. Let P_i denote the set of prerequisites of learning objective $i \in U$. If $j \in P_i$, that is if learning objective j is a prerequisite to learning objective i, it is possible to assume that if i is assessed, then implicitly j is assessed too. In such a case, the Full covering axiom admits a natural variant, which imposes, for each course $N \in 2^U$, that

$$\bigcup_{i \in \bigcup\limits_{S \in f(N)} S} \big(P_i \cup \{i\}\big) = N.$$

These extensions are left for future works.

Appendix

Proof. (Proposition 1) [If part] For any partition P of U, denote by f^P the unique examination function such that, for each $N \in 2^U$, $f^P(N) = \{(S \cap N) \neq \varnothing : S \in P\}$. Fix some course N. By definition of a partition, for each partition P of U and each N, $f^P(N)$ is a partition of N. Hence, any learning objective $i \in N$ belongs to one and only one exercise in $f^P(N)$. This implies that $f^P(N)$ covers N, that for each exercise $S \in f^P(N)$,

$$\bigcup_{T \in f^P(N) \setminus \{S\}} T = N \setminus S \neq N,$$

and that for each learning objective $i \in N$, $|\{S \in f^P(N) : i \in S\}| = 1$, which proves that f^P satisfies **FC**, **NR** and **B**. Regarding **AC**, for any learning objective $i \in N$, we have

$$f^P(N \setminus \{i\}) = \{(S \cap N \setminus \{i\}) \neq \varnothing : S \in P\} = \{(S \setminus \{i\} \cap N) \neq \varnothing : S \in P\} = \{(S \setminus \{i\}) \neq \varnothing : S \in f^P(N)\},$$

and so f^P satisfies **AC**.

[Only if part] Let f be any examination function satisfying **FC**, **NR**, **AC** and **B**. We split the demonstration in two steps.

Step 1. We first show that for each $i \in U$, it holds that $|\{S \in f(U) : i \in S\}| = 1$.

Note that **B** implies that there exists an integer $m \geq 0$ such that, for each $i \in U$, $|\{S \in f(U) : i \in S\}| = m$, and **FC** imposes that $m \geq 1$.

By way of contradiction, suppose that there exists $i \in U$ and two different exercises $S, T \in f(U)$ such that $i \in S \cap T$. Denote by $S\Delta T$ the symmetric difference between S and T, defined by $S\Delta T = (S \backslash T) \cup (T \backslash S)$. Let $j \in S\Delta T$. Without loss of generality, one may suppose that $j \in S$, so that $j \notin T$. Now **AC** implies that both $\{i,j\}$ and $\{i\}$ are in $f(\{i,j\})$, because $\{i,j\} \subseteq S$ while $i \in T$ and $j \notin T$. Then $\bigcup_{R \in f(\{i,j\}) \backslash \{i\}} R = \{i,j\}$ which violates **NR**. Hence $m \leq 1$.

Step 2. The preceding step ensures that there exists a unique partition P of U such that $f(U) = f^P(U)$. Now, for any $N \in 2^U$, we have

$$\begin{array}{rl} & f(N) \\ = & f(U \backslash (U \backslash N)) \\ = & \{S \backslash (U \backslash N) \neq \varnothing, S \in f(U)\} \\ \overset{\text{Step 1}}{=} & \{S \backslash (U \backslash N) \neq \varnothing, S \in f^P(U)\} \\ = & \{S \cap N \neq \varnothing, S \in f^P(U)\} \\ = & \{S \cap N \neq \varnothing, S \in P\} \\ = & f^P(N), \end{array}$$

which completes the proof. ■

The logical independence of the axioms invoked in Proposition 1 is demonstrated by the following examination function:

- The examination function f which assigns to each $N \subseteq U$ the empty exam $f(N) = \varnothing$ satisfies **NR**, **AC** and **B** but violates **FC**.
- Choose any integer k such that $2 \leq k \leq u$. The examination function f which assigns to each $N \subseteq U$ the exam $f(N) = \{S \in 2^N : 1 \leq s \leq k\}$ satisfies **FC**, **AC** and **B** but violates **NR**.
- For each $N \subseteq U$, let $i_N = \min_{i \in N} i$ be the learning objective with the "smallest" index in N. The examination function f which assigns to each $N \subseteq U$ the exam$f(N) = \{i_N\} \cup \{N \backslash \{i_N\}\}$ satisfies **FC**, **NR** and **B** but violates **AC**. To see this, let $N = \{1,2,3\}$. Then, $f(\{1,2,3\}) = \{\{1\},\{2,3\}\}$ and $f(\{2,3\}) = \{\{2\},\{3\}\}$ whereas **AC** applied to $f(\{1,2,3\})$ and $\{2,3\}$ yields $f(\{2,3\}) = \{\{2,3\}\}$.
- Choose $i \in U$ and a partition P of $U \backslash \{i\}$. Consider the set $P_i = \{S \cup \{i\} : S \in P\}$. The examination function f which assigns to each $N \subseteq U$ the exam $f(N) = \{(S \cap N) \neq \varnothing : S \in P_i\}$ satisfies **FC**, **NR** and **AC** but violates **B**.

Proof. (Proposition 2) [If part] The examination function f which assigns to each N its finest partition satisfies **FC** and **NR** by Proposition 1. It also obviously satisfies **D**.

[Only if part] Let f be any examination function satisfying **FC**, **NR** and **D**. By **FC**, it is enough to show that $s = 1$ for all $N \subseteq U$ and all $S \in f(N)$. By contradiction, assume that there are nonempty course $N \subseteq U$ and exercise $S \in f(N)$ such that $s > 1$. Choose any learning objective $i \in S$. By **D**, there must exist another exercise $T \in f(N)$ such that $T \ni i$ and $T \neq S$. By **FC** $\bigcup_{R \in f(N)} R = N$. So, if $T = \{i\}$, we get $\bigcup_{R \in f(N) \setminus \{T\}} R = N$ since $i \in S \in f(N)$. This contradicts **NR**. Hence, we get that $t \geq 2$, which means that learning objective i is assessed in at least two exercises of size at least 2. Since S was an arbitrary exercise of size at least 2 and i an arbitrary learning objective in this exercise, we conclude that any learning objective assessed in any exercise of size at least 2 is also assessed in another exercise of size at least 2. As a consequence, for the initially chosen exercise S, we get

$$\bigcup_{R \in f(N) \setminus \{S\}} R = \bigcup_{R \in f(N)} R \overset{\mathbf{FC}}{=} N,$$

which contradicts the fact that f satisfies **NR**. This completes the proof. ∎

The logical independence of the axioms invoked in Proposition 1 is demonstrated by the following examination function:

- The examination function f which assigns to each $N \subseteq U$ the empty exam $f(N) = \varnothing$ satisfies **NR** and **D** but violates **FC**.
- The examination function f which assigns to each $N \subseteq U$ the exam $f(N) = C(n, N)$ satisfies **FC** and **NR** but violates **D**.
- The examination function f which assigns to each $N \subseteq U$ the exam $f(N) = C(2, N)$ satisfies **FC** and **D** but violates **NR**.

Proof. (Proposition 3) Let f be an examination function that satisfies **N** and **SC**. We shall only prove part *(i)* since parts *(ii)*, *(iii)* and *(iv)* immediately follow from *(i)*. So, by contradiction, suppose that there exists $N \subseteq U$ and $i \in N$ such that $\{i\} \notin f(N)$. By $n - 1$ successive applications of **SC** to the learning objective $j \in N \setminus \{i\}$, we obtain that $f(\{i\}) = \varnothing$, a contradiction with the fact that f satisfies **N**. ∎

Proof. (Proposition 4) [If part] For each $Q \subseteq \{1, \ldots, u\}$ with $Q \ni 1$, define by f^Q as the examination function such that, for each $N \in 2^U$, $f^Q(N) = \{C(q, N), q \in Q\}$. Then f^Q satisfies **N** and, for each $N \in 2^U$, $q \leq n$ and $i, j \in N$, $|\{S \in C(q, N) : S \ni i\}| = |\{S \in C(q, N) : S \ni j\}|$ and

$$\sum_{S \in C(q,N) : S \ni i} 1/s = \sum_{S \in C(q,N) : S \ni j} 1/s$$

so that f^Q satisfies **B** and **W**. Lastly, for each $N \in 2^U$, $q < n$ and $i \in N$, $C(q, N \setminus \{i\}) = \{S \in C(q, N) : S \not\ni i\}$ so that f^Q satisfies **SC**.

[Only if part] Consider any examination function f that satisfies one of the two triple of axioms, we show that there is some $Q \subseteq \{1, \ldots, u\}$ with $Q \ni 1$ such that $f = f^Q$. We split the rest of the demonstration in three steps.

STEP 1. Given $S \in 2^U$ and $i \notin S$, then $S \cup \{i\} \in f(U)$ implies that $S \cup \{j\} \in f(U)$ for any $j \in U \backslash S$.

By way of contradiction, consider the triples $(S, i, j) \in 2^U \times U \times U$ such that $i, j \notin S$, $S \cup \{i\} \in f(U)$ and $S \cup \{j\} \notin f(U)$. Pick one such triple so that S has the smallest size. Note that, by iterating **SC** on $U \backslash (S \cup \{i, j\})$, $S \cup \{i\} \in f(S \cup \{i, j\})$ and $S \cup \{j\} \notin f(S \cup \{i, j\})$. Moreover, by minimality of S, there is neither strict subset $T \subset S \cup \{i, j\}$ such that $T \cup \{i\} \in f(S \cup \{i, j\})$ and $T \cup \{j\} \notin f(S \cup \{i, j\})$, nor strict subset $T' \subset S \cup \{i, j\}$ such that $T' \cup \{j\} \in f(S \cup \{i, j\})$ and $T' \cup \{i\} \notin f(S \cup \{i, j\})$.

Then one has the following partition $\{R \in f(S \cup \{i, j\}), R \ni i\} = \{R \in f(S \cup \{i, j\}), R \supseteq \{i, j\}\} \cup \{R \in f(S \cup \{i, j\}), R \ni i \text{ and } R \not\ni j \text{ and } R \cup \{j\} \backslash \{i\} \in f(S \cup \{i, j\})\} \cup \{S \cup \{i\}\}$. Hence, any sum on $\{R \in f(S \cup \{i, j\}), R \ni i\}$ only involving the cardinals has exactly one more term than the sum on $\{R \in f(S \cup \{i, j\}), R \ni j\}$. This shows that f can neither satisfy **B** nor **W**.

STEP 2. There is a unique $Q \subseteq \{1, \ldots, u\}$ such that $f(U) = f^Q(U)$.

By iterating STEP 1, if $S \in f(U)$, then any $S' \in 2^U$ such that $s' = s$ also belongs to $f(U)$. Finally, $f(U) = \{C(s, U), S \in f(U)\}$ which can be uniquely written as $f^Q(U)$. By Proposition 3 (i), $1 \in Q$.

STEP 3. We finally show that $f = f^Q$ on 2^U.

For each $S \subseteq N \in 2^U$, by iterating **SC** on $U \backslash N$, $S \in f(N)$ if and only if $S \in f(U) = f^Q(U)$, and $S \in f(U) = f^Q(U)$ if and only if $S \in f^Q(N)$. The proof is complete. ■

The logical independence of the axioms invoked in Proposition 4 is demonstrated by the following examination function:

- The examination function f which assigns to each $N \subseteq U$ the empty exam $f(N) = \varnothing$ satisfies **B**, **W** and **SC** but violates **N**.
- Fix some $S \in 2^U$ such that $1 < s < u$. The examination function f which assigns to each $N \subseteq U$ the exam $f(N) = C(1, N) \cup \{S\}$ if $S \subseteq N$ and $f(N) = C(1, N)$ otherwise satisfies **N** and **SC** but violates both **B** and **W**.
- For each $N \in 2^U$, choose $Q_N \subseteq \{1, \ldots, n\}$ such that $1 \in Q_N$, and assume that $Q_N \nsubseteq Q_{N \cup \{i\}}$ for some $i \in U$ and some nonempty $N \subseteq U \backslash \{i\}$. The examination function f which assigns to each $N \subseteq U$ the exam $f(N) = \{C(q, N) : q \in Q_N\}$ satisfies **N**, **B** and **W** but violates **SC**.

Proof. (Proposition 5) For any ordering π on U, denote by f^π the examination function such that, for each $N \in 2^U$, $f^\pi(N) = \{P_{i,\pi} \cap N, i \in N\}$. It is clear that f^π satisfies both **GC** and **AC**. This examination function is also minimal with respect to **GC** and **AC** since, for each course $N \in 2^U$, **GC** imposes at least n exercises and $|f^\pi(N)| = n$.

It remains to show that if an examination function f satisfies **GC** and **AC** and is minimal with respect to these two axioms, then there is some ordering π on U such that $f = f^\pi$. So pick any such examination function f. For any ordering π, we already know that f^π satisfies **GC** and **AC** and is also minimal with respect to these axioms. Therefore, for any $N \in 2^U$, $f(N)$ must consist of n exercises of size from 1 to n. Note also that $N \in f(N)$. We split the rest of the demonstration in two steps.

Step 1. For all $N \in 2^U$, There is a unique learning objective $i \in N$ such that $f(N) = f(N\backslash\{i\}) \cup \{N\}$.

Uniqueness comes from the fact that $f(N) = f(N\backslash\{i\}) \cup \{N\} = f(N\backslash\{j\}) \cup \{N\}$ implies that $N\backslash\{i\}$ and $N\backslash\{j\}$ are in $f(N)$, which contradicts the minimality with respects to **GC** and **AC**. Moreover, if $N \neq \varnothing$, $f(N)$ contains an exercise of size $n-1$ included in N. Hence, there is a learning objective $i \in N$ such that $N\backslash\{i\} \in f(N)$. By contradiction, assume that there is other exercise $S \in f(N)$ so that $S \ni i$ and $s < n-1$. Consider S_0 the smallest of such exercise. If $s_0 > 1$, then there is an exercise $T \in f(N)$ with $t = s_0 - 1$ which does not assess i. Then $T \in f(N\backslash\{i\})$ and so is $S_0\backslash\{i\}$, which contradicts the minimality of f with respects to **GC** and **AC** as there are two exercises of size $s_0 - 1$ in $f(N\backslash\{i\})$. This implies that $s_0 = 1$ and $S_0 = \{i\}$. Hence $N, N\backslash\{i\}$ and $\{i\}$ are three of the n exercises in $f(N)$. Now, by **AC**, $f(N\backslash\{i\}) = \{(S\backslash\{i\}) \neq \varnothing, S \in f(N)\}$ consists of at most $n-2$ exercises, which contradicts **GC**.

Step 2. There is some ordering π on U such that $f(U) = f^\pi(U)$.

By recursive application of Step 1 starting from U, there exists a sequence $(k_1, \ldots, k_{u-1})$ of learning objectives of U so that $f(U) = \{U, U\backslash\{k_1\}, U\backslash\{k_1, k_2\}, \ldots, U\backslash\{k_1, \ldots, k_{u-1}\}\}$. The last exercise $U\backslash\{k_1, \ldots, k_{u-1}\}$ defines a unique learning objective $k_u \in U$. Then, beginning with the end, there is a unique ordering π on U, defined by $\pi^{-1}(i) = k_{u-i+1}$ for any $1 \leq i \leq u$ and

$$f(U) = \{P_{i,\pi}, i \in U\} = f^\pi(U). \tag{1}$$

Now, for any $N \in 2^U$, we have

$$\begin{aligned}
& f(N) \\
= \;& f(U\backslash(U\backslash N)) \\
= \;& \{S\backslash(U\backslash N) \neq \varnothing, S \in f(U)\} \\
\overset{(1)}{=} \;& \{S\backslash(U\backslash N) \neq \varnothing, S \in f^\pi(U)\} \\
= \;& \{S \cap N, S \in f^\pi(U)\} \\
= \;& \{P_{i,\pi} \cap N, i \in U\} \\
= \;& \{P_{i,\pi} \cap N, i \in N\} \\
= \;& f^\pi(N),
\end{aligned}$$

which completes the proof. ■

The logical independence of the axioms invoked in Proposition 5 is demonstrated by the following examination function:

- The examination function f which assigns to each $N \subseteq U$ the empty exam $f(N) = \varnothing$ satisfies **AC** and is minimal with respect to this axiom but violates **GC**.

- For each nonempty $N \in 2^U$, fix an ordering π_N in such a way that for some $N, M \in 2^U$ with $M \subsetneq N$, the restriction of π_N to M does not coincides with π_M. The examination function f which assigns to each $N \subseteq U$ the exam $f(N) = \{P_{i,\pi_N}, i \in N\}$ satisfies **GC** and is minimal with respect to this axiom but violates **AC**.
- The examination function f which assigns to each $N \subseteq U$ the exam $f(N) = 2^N \backslash \varnothing$ satisfies **AC** and **GC** but is not minimal with respect to these two axioms.

References

Airasian, P. W., Miranda, H., 2002. The role of assessment in the revised taxonomy. Theory into Practice 41, 249–254.

Bilbao, J. M., 2000. Cooperative games on combinatorial structures. Kluwer Academic Publishers, Boston.

Bloom, B. S., 1956. Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain, 2nd Edition. Addison-Wesley Longman Ltd.

Bondareva, O. N., 1963. Some applications of linear programming methods to the theory of cooperative games. Problemy kibernetiki (in russian) 10, 119–139.

Harden, R. M., 2001. AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning. Medical Teacher 23, 123–137.

Hart, S., Kurz, M., 1983. Endogenous formation of coalitions. Econometrica 51, 1047–1064.

Peleg, B., 1965. An inductive method for constructing minimal balanced collections of finite sets. Naval Research Logistics Quarterly 12, 155–162.

Shapley, L. S., 1967. On balanced sets and cores. Naval Research Logistics Quarterly 14, 453–460.

Thomson, W., 2019. Consistent Allocation Rules. Cambridge University Press.

van den Brink, R., 2012. On hierarchies and communication. Social Choice and Welfare 39, 721–735.

Wallis, W. D., 2007. Introduction to Combinatorial Designs, 2nd Edition. Chapman and Hall/CRC.

Walvoord, B. E., 2010. Assessment Clear and Simple: A Practical Guide for Institutions, Departments, and General Education, 2nd Edition. Jossey-Bass.

Walvoord, B. E., Johnson Anderson, V., 2009. Effective Grading, 2nd Edition. Jossey-Bass.

Zapechelnyuk, A., 2015. An axiomatization of multiple-choice test scoring. Economics Letters 132, 24–27.