Physica A 436 (2015) 418–429

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

The detection of cheating in multiple choice examinations

Peter Richmond [a], Bertrand M. Roehner [b,*]

[a] School of Physics, Trinity College Dublin, Ireland

[b] Institute for Theoretical and High Energy Physics (LPTHE), University Pierre and Marie Curie, Paris, France

HIGHLIGHTS

- Identification of connected answers.
- Two methods are proposed.
- Simulations show role of parameters.

ARTICLE INFO

Article history:
Received 8 February 2015
Received in revised form 3 April 2015
Available online 15 May 2015

Keywords:
Examination
Cheating
Detection
Binomial trial
Multinomial trial
Dependent variables

ABSTRACT

Cheating in examinations is acknowledged by an increasing number of organizations to be widespread. We examine two different approaches to assess their effectiveness at detecting anomalous results, suggestive of collusion, using data taken from a number of multiple-choice examinations organized by the UK Radio Communication Foundation. Analysis of student pair overlaps of correct answers is shown to give results consistent with more orthodox statistical correlations for which confidence limits as opposed to the less familiar "Bonferroni method" can be used. A simulation approach is also developed which confirms the interpretation of the empirical approach.

1. Background

Cheating in examinations is now acknowledged by an increasing number of academic institutions to be widespread. Wesolowsky [1] notes one 2005 US study [2] which reported that from 10,000 faculty, 44% were aware of student cheating in the previous 3 years yet never reported the fact. He also points to a second study by the Josephson Institute of Ethics also in USA. This surveyed 43,000 high school students in October 2010, which found that more than one half admitted cheating on test during past 12 months and one third admitted cheating more than twice. Cizek [3] as a result of a survey of studies prior to 1999 concluded that "cheating is rampant". Most research since suggests cheating is more widespread than is usually believed. In the UK cheating has been reported to occur during medical examinations [4]. But other specialisms are not exempt. Nor is it confined to schools and universities. Many professionals, aspiring car drivers, nuclear missile control supervisors and amateur radio operators are all required to take examinations in order to gain an appropriate licence. It is also not unknown for the teacher to collude with examinees. This was illustrated by a recent article in Time Magazine [5] which reported the indictment of a college Principal and four teachers for conspiracy and provision of examination answers to students before the examination. Multiple-choice examinations form a significant element of the testing process in all these cases. Here we report the development and implementation of a statistical method aimed at detecting cheating in

* Corresponding author.
E-mail addresses: peter_richmond@ymail.com (P. Richmond), roehner@lpthe.jussieu.fr (B.M. Roehner).

http://dx.doi.org/10.1016/j.physa.2015.05.040

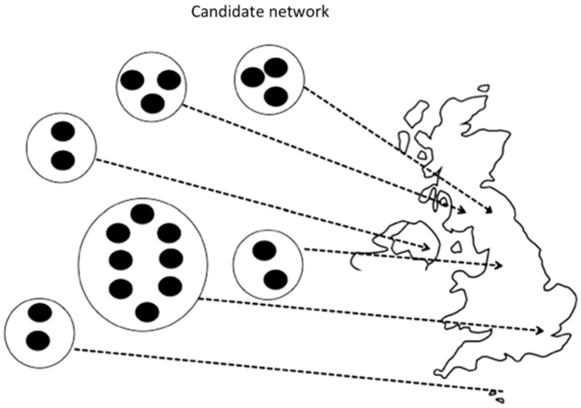

Fig. 1. Schematic illustration of the organization of examinations across the UK with relatively small numbers of students taking the examination at different centres simultaneously.
Source: Office of Communications Regulation (OFCOM) and Radio Communications Foundation (RCF).

multiple choice amateur radio examinations. Our empirical study is complemented by a set of simulations, which offer additional insight into the method and outcomes.

The next section outlines the way amateur radio examinations in the UK are administered and managed together with initial approaches to cheating detection. A statistical approach to detection is then described together with some output for three different examinations that illustrate the potential of the method. Finally we end with thoughts and conclusions.

Amateur radio has its origin in the early studies by Marconi and other physicists of electromagnetic wave propagation at the end of the 19th and beginning of the 20th centuries. As a hobby it began to grow substantially after world war one. It has been estimated that around two million people world wide are regularly involved with amateur radio. In the UK for the past few years new entrants number about 1000 annually. To participate, entrants across the world must obtain a licence. Most advanced countries issue this following success in a national examination. In the UK, the government office of communications regulation (OFCOM) confers responsibility for the examination to the Radio Communications Foundation (RCF) an independent charity established to support people and projects where radio communication, and amateur radio in particular, is the theme. The examination itself provides for three levels of licence: Foundation, Intermediate and Advanced. Here we are concerned with the Advanced examination which at the moment consists of 62 multiple choice questions that test knowledge of various aspects of radio including the licence conditions, basic electronics, radio and transmitter architecture, antennae and electromagnetic wave propagation, testing methods and safety. Each question is constructed with the correct answer and three distractors. The examination is held simultaneously at a number of approved centres located across the UK. The number of candidates at any centre varies depending on demand (see schematic shown in Fig. 1). As an example, in 2008 63 candidates sat for an advanced examination across 25 centres. The average number of candidates per centre varies from 1 to 6.

Locations are vetted and the conduct of the examination follows the traditional pattern found in most schools and academic establishments. Invigilators are required to be named and deemed suitable by the RFC prior to the examination. However unlike schools, the invigilators are volunteers and unpaid. Optical mark sheets record the answers and for each candidate one has therefore a 62 character string, viz: *ABDC . . . BCADC* plus two additional identifiers for the candidate and the particular centre in which the candidate sat the examination.

It can be noted that for confidentiality reasons all identification code numbers of candidates and examination centres have been removed from the results given in the paper.

2. Analysis of empirical data

In studying the data we follow a method similar to that proposed by McManus et al. [4] who examined the overlap, A_{ij} of correct answers for each and every pair of candidates in their cohort as a function of the geometric mean of the scores, A_i and A_j for the pair of students. For the definition of overlap and related notions, see Table 1. In our case it is useful to not only consider the total network of all student pairs but also the sub-network formed by pairs of students which sat the examination in each separate centre. A feature of the scatter plot such as shown in Fig. 2 is that each data point falls entirely inside the limited area defined by the dotted triangle. This may be understood when one realizes that the maximum value of the overlap occurs when the scores for each student are equal; the minimum value of the overlap depends on the values scored. If the students score less than 50%, the minimum possible value of the overlap is zero. For students scoring over 50% the minimum overlap increases linearly from zero to the maximum value when each student scores 100%.

Assuming no contact by phone, internet or other wireless means during the examination, we may expect the total overlap scores for the total student network to be completely uncorrelated since it is dominated by pairs from different examination