The Innovation Journal: The Public Sector Innovation Journal, Volume 15(1), 2010, Article 1.

Some forecasts of the diffusion of e-assessment using a model

Andrew Boyle
Head of Assessment Research
Office of the Qualifications and Examinations Regulator (Ofqual)
Coventry, England

Andrew.Boyle@ofqual.gov.uk

Some forecasts of the diffusion of e-assessment using a model

Andrew Boyle

ABSTRACT

This article concerns the use of e-assessment in General Certificate of Secondary Education (GCSE) qualifications in England. It explains why the diffusion of e-assessment merits study. It reviews approaches to forecasting the diffusion of e-assessment in GCSEs and justifies the choice of an aggregate diffusion model (the Bass model) for this purpose.

The article discusses approaches to the parameterisation of the Bass model and proposes three sets of parameters for the model. The three disparate forecasts of e-assessment diffusion result. The most accurate of the three forecasts will become clearer as more data of e-assessment adoption are collected.

The article suggests extensions and generalisations to the core Bass model, which may provide an enhanced understanding of e-assessment adoption for GCSEs.

It is informative to think about e-assessment diffusion with reference to schools' decisions to adopt the new technology. First, forecasts should be made, using the core model, and then possible extensions should be proposed. Importing ideas from marketing research into the sphere of exams and qualifications is a useful example of how work from the wider social sciences can inform assessment research. Diffusion forecasting could have a role in understanding other qualifications and examinations issues.

Key Words: Bass model; education; national examinations and qualifications systems; electronic assessment (e-assessment)

Background

The English qualifications system

England is a constituent country of the United Kingdom (UK), itself a Member State of the European Union. This article relates to educational qualifications in England other than those awarded by Higher Education Institutions.

Young people in England typically study towards separate qualifications for different subjects. Although not unique internationally, this practice is unusual. In most countries students study towards a unitary school-leaving certificate or diploma (Le Métais, 2002, p. 24). Typically, students in England study for eight to ten General Certificate of Secondary Education (GCSE) qualifications at the age of about 16 – currently the end of compulsory education. Thereafter, many students spend the next two years of post-compulsory study towards General Certificate of Education (GCE) Advanced level (A level) qualifications; these contain an Advanced Subsidiary (AS) qualification that can be awarded separately after one year of study.

GCSEs and GCE A levels can be classified as 'general' or 'academic' qualifications. However, other types of qualifications exist. Such types include: Vocationally Related Qualifications (VRQs) or Other General (OG) qualifications, such as graded music or language qualifications. At the time of writing, there are many new developments in qualifications. Several 'composite' qualifications (i.e. where one overarching qualification that 'contains' several elements, including other qualifications) are being developed. The suite of Diplomas is the most prominent of these (DCSF, 2007).

The structure of qualifications provision in England is also different from that in many other nations. In many nations it is a Ministry of Education, or a single technical agency that develops, administers, marks and awards (or 'certifies') school qualifications (see Boyle, 2008). In England, qualifications providers (awarding bodies – ABs) are independent organisations, and schools are free to choose the AB that suits their needs to provide qualifications for their students. Several commentators have characterised this as a 'qualifications market' and have started to describe the market’s features (PWC, 2005; QCA, 2008). However, others have suggested that some behaviours of organisations and individuals may not be consistent with the existence of a conventional competitive market (Europe Economics, 2008). There are three ABs in England and one each based in Wales and Northern Ireland that are entitled to award GCSEs and GCEs[i].

Qualifications ABs are regulated by the independent Office of the Qualifications and Examinations Regulator (Ofqual). Ofqual, working with its sister regulators in Wales and Northern Ireland, accredits qualifications, and maintains databases of accredited qualifications. The regulators are also able to determine whether a qualification is run using e-assessment, and to request that ABs share information, for instance on the number of centres that are using e-assessment.

e-assessment

For the purposes of this article, electronic assessment (e-assessment) – also known as computer-based or computer-assisted assessment – includes examinations or tests run on computer, but also other technology-supported assessment methods such as e-portfolios. The term is defined here to include only those instances where the examinee interacts with the assessment material via a computer (be that a desktop, laptop, handheld, etc.). (For other approaches to defining e-assessment, see: JISC, 2006 and QCA et al, 2007.)

The stimulus for this research

Imagine a child born in 1999. She will grow up with Information and Communication Technologies (ICTs) as an integral part of her life. From an early age, she has interacted with learning and entertainment content through websites, seen mobile phones and text messaging as a normal part of life.

If this child is typical, she will take her GCSEs in 2015. She is a natural user of ICT, and much of her learning will have been supported by technology. But what will happen when she sits her examinations? Will she have to put down the mouse and pick up a pen? Anecdotally, even at the time of writing (2008) some young people have to 'learn how to handwrite' (quickly, neatly) just to sit exams. If exams remain solely or mainly on paper, examination candidates will increasingly be denied tools that they have always used when composing text, solving mathematical problems or finding out information. When this child takes her exams in 2015, will they still be on paper, or will she use e-assessment?

Next, consider an awarding body. It has a portfolio of qualifications that it owns and runs. Evidence from the 2008 'qualifications market report' would suggest that the AB's portfolio of qualifications was increasing year on year (QCA, 2008, p. 14); the AB may be developing new composite qualifications (*ibid.* at p. 17) – a process that could well entail a step into the dark, the commitment of considerable resource and significant risk. If it is a typical large awarding body, the AB will be taking early steps in the use of technology to assist the examinations process. This might include the substantial use of e-marking, as well as some use of ICTs for other processes such as awarding or results publishing (*ibid.* at pp. 33ff.). Finally, this typical large AB would have a range of qualifications that included e-assessment. In 2008 these would typically be VRQs or OGs – although the AB would be likely to have one or two GCSEs or GCEs containing some e-assessment; these perhaps being in a pilot phase or operating with a very small cohort.

Given the range of significant demands on the typical large AB described above, they may well wish to know how quickly e-assessment usage will grow; and at what point in time most of their large cohort qualifications will be run on-screen. Further, AB managers and leaders may wish to know whether they have any levers to accelerate diffusion of e-assessment. Also, they may need to know whether the nature of diffusion will give rise to any substantive or resource implications, so that they can plan how to structure their business.

Finally, in this section, consider the qualifications regulator. It has made statements supporting the diffusion of e-assessment (Boston, 2004), and has introduced a regulatory approach that it believes to be congruent with that support for innovation and to embody the UK government's principles for better regulation (QCA et al, 2007). The regulator will want to know whether its support for innovation has been successful in fostering the wider use of e-assessment. Indeed, the regulator will need to establish how to quantify what 'success' means in terms of awarding bodies providing e-assessment and schools and colleges offering it to their students. How long should the regulator expect it to take for e-assessment to be commonplace? Such a question could cut both ways; on the one had, the regulator might feel that e-assessment was being introduced too slowly, and might therefore wish to 'stimulate the market'. Alternatively – if e-assessment were to take off very quickly

– this might provide a range of issues (e.g. around technical infrastructures in schools, relating to the standards that the qualifications embodied, etc.) about which the regulator had to have a position.

These three contrasting perspectives on the issue of e-assessment provide the stimulus for this research.

Research aim

The aim of the research described in this paper is to derive and evaluate some credible forecasts of the diffusion of e-assessment as an assessment method in qualifications used in England. Several quantitative aggregate forecasts of the diffusion of e-assessment in the specified qualifications will be derived.

It will not seek (as a primary finding from the research) to explain *why* diffusion of e-assessment is or is not likely to happen; however, it will seek to discuss substantive implications of the forecasted diffusion of e-assessment.

In order to put a practical limit on the research (and also in order to be able to track the accuracy of the forecast in future work), it will relate only to the diffusion of e-assessment in GCSE qualifications that are accredited by UK qualifications regulators.

This is not the first time that a researcher has predicted the future diffusion of e-assessment. Other reputable commentators such as Kingdon (2005) and Ripley (2006) have done so. However, those two 'predictions' were not carried out according to an explicit methodology and were as much commentaries and arguments about current and future policy directions as they were forecasts (which is not a bad thing). However, the current researcher believes that this is the first attempt to use a well-known forecasting model to predict the diffusion of e-assessment. It is hoped that, as such, this will be a contribution to the research literature, and a starting point for the systematic investigation of e-assessment diffusion.

Review of research literature

Definitions of key terms

'Innovation' has many potential meanings. The first distinction to be made is between the non-count noun – innovation as a concept, and the tangible, countable thing 'an innovation'. Whilst the current research certainly has implications for innovation as a concept (e.g. whether qualifications provision and the wider education sector is innovative or not), it is centrally concerned with e-assessment as a tangible, countable innovation.

Rogers (2003, p. 36) defines an innovation as:

> … an idea, practice, or object perceived as new by an individual or other unit of adoption

This definition does not suggest that the innovation is necessarily a good thing. For instance, one can model the diffusion of e-assessment without considering whether this diffusion is likely to make examinations more or less valid.

The term 'diffusion' needs to be defined. Rogers (2003, p. 5) does so as follows:

> *Diffusion* is the process in which an innovation is communicated through certain channels over time among members of a social system. It is a special type of communication, in that the messages are concerned with new ideas.

Rogers also writes about diffusion as:

> a kind of *social change*, defined as the process by which alteration occurs in the structure and function of a social system. When new ideas are invented, diffused and adopted or rejected, leading to certain consequences, social change occurs. (Rogers, 2003, p. 6)

The term 'adoption' is closely associated with diffusion. Some writers use the two terms more or less synonymously (e.g. Meade & Islam, 1998; Van den Bulte, 2000, p. 367). However, it is important to understand the term 'adoption' in its own right. Firstly, one may contrast it with 'diffusion'. The first juxtaposition is that 'diffusion' refers to the general (aggregate-level) process by which an idea, practice or object is communicated in a social system, whilst adoption is better used to refer to the decisions that individuals in that social system make each time that they consider taking up an innovation (Rangaswamy & Gupta, 2000, p. 76; Hauser, Tellis & Griffin, 2006, p. 689).

Further light can be thrown on the distinction between 'diffusion' and 'adoption' by considering the academic disciplines from which the terms stem. Rogers was originally a rural sociologist (Rogers, 2003, p. xvii), but developed his work to posit a generic process of communication in social systems. Bass (1969) derived considerable theoretical underpinning from Rogers' work. Bass, in turn, is the father of a strand of research that explores, extends and generalises a quantitative aggregate model for the diffusion of consumer durable products[ii], which is widely used in marketing science. Rogers' work was widely applied from the outset. In contrast, aggregate diffusion models such as Bass's have a narrower basis – Bass's original model was concerned with initial, single purchases of consumer durable ('white') goods (Bass, 1969, p. 216)[iii]. Thus, in the context of such aggregate diffusion models an individual 'adoption' is sometimes used as a synonym for a 'purchase' (Chandrasekaran & Tellis, 2008, p. 48) and the overall number of adoptions is often considered to be equal to the volume of sales for the product category concerned (Mahajan, Muller & Bass, 1990, p. 2).

Finally, in defining adoption, one can make explicit a point which is strongly implied several times above. The adopter is the person or organisation that purchases a new product or service (or that decides to take up an innovation). For the purposes of this definition, adoption does not refer to the decision of a product or service provider to make an innovation available. Therefore, in the context of GCSE qualifications in England, the adopting unit will be schools that decide to start using a GCSE qualification that contains e-assessment. Of course, one might expect schools' decisions to adopt e-assessed qualifications to be influenced by the extent to which awarding bodies provide such qualifications. Also, exam candidates' opinions might also affect schools' decisions to take up e-assessment. However, the models discussed in this paper are essentially 'demand side'; assuming that the best understanding of adoption of e-assessment can be gained from models that describe potential 'purchasers' and factors that affect them. Whilst that assumption could be disputed, the model used in this research has been shown across many applications to be a powerful empirical generalisation that fits data from many different innovations and contexts (Bass, 1995).

Aggregate diffusion models

The broad choice of model type

Literature describing the origins and application of diffusion models is vast and cuts across several disciplinary boundaries (Geroski, 2000, p. 604). Any brief summary of key trends in that sprawling body of knowledge inevitably provides a simplified picture of that large research field. Nonetheless, it is necessary in this article to extract the main features of diffusion models.

A key differentiator between families of diffusion models concerns what the key explanator of the diffusion rate is taken to be. The explanator could be either features of communication within an essentially homogenous population, or the heterogeneity of members of the population with respect to some characteristic (other than communication) that influences population members to adopt an innovation (Geroski, 2000, p. 610). A wide range of matters can be considered in innovations studies that assume a heterogeneous population. Studies can model the impact of differences between population members in respect of (for instance) firm size, technology expectations and the cost of switching to a new technology (Geroski, 2000, pp. 611 – 614). Other studies can model changes within individuals' attitudes to innovations over time (Chatterjee and Eliashberg, 1990).

The choice of this research is to model the diffusion of e-assessment subject to the assumption of an essentially homogenous population of potential adopters. There are sound reasons this preference, which include:

- The qualifications that candidates enter for are chosen by schools and colleges (including any choices as to whether to enter for an e-assessed qualification). Those institutions are, of course, very diverse in many ways. However, the professional staff in institutions who contribute to choices about qualifications (who might include: school or college senior leaders, teachers, network managers and ICT technicians, and exams officers) collectively constitute a more homogenous population than would be the case in many investigations of the adoption of an innovation, where the target group is literally the entire population of a country. For example, schools have functions and structures that are specified by law and regulated by bodies such as schools' governing bodies and the Office for Standards in Education, Children's Services and Skills (Ofsted), amongst others. Professionals in schools must have specified qualifications (the qualifications held by members of the general population are naturally more heterogeneous), and their working tasks are also prescribed – for example by the national curriculum or by other legal obligations on schools. As such, it makes sense to use a model that assumes a homogenous population to predict the adoption of e-assessment by schools.
- Although heterogeneity models to explain diffusion can study many variables, one of the chief issues that this strand of research has addressed is that of price; including the relationship between the price of an innovation and the distribution of incomes within a (potentially) adopting population (Geroski, 2000, p. 620). However, one of the key reasons that Europe Economics believed the English qualifications market to be at best an atypical market was that price is not typically a major determiner in decisions about purchasing entry to a qualification (Europe Economics, 2008, p. 15).

Therefore this research will be based around a diffusion model that takes aspects of communication functioning within an otherwise essentially homogenous population to be the main explicators of diffusion.

Choosing a specific communication-based model

Within the family of models that take communication as the main explanator of diffusion, there are several competing models. As in the previous section, simplifying this complex area will necessarily omit some aspects of the research knowledge base. Nonetheless, a summary of main strands is consistent with the aims of this article.

Teng, Grover & Güttler's (2002, pp. 14 – 15) account of innovation diffusion models includes three types of models, which are described below:

1. *external influence model*; assumes that all influence on members of a social system to adopt an innovation comes from external communications to that social system. External influences are often equated with mass media communications, but conceptually the term refers to communication via vertical, centralised, structured or formal channels (Mahajan & Peterson, 1985, p. 15). The cumulative growth of adopters under the external influence model is described by a decaying or modified exponential curve; that is, over time, the cumulative number of adopters increases, but at a constantly decreasing rate (*ibid.* at p. 16).
2. *internal influence model*; assumes that all influence on social system members to adopt an innovation comes from other members of that social system, and not from external influences. This model typically yields a logistic ('s-shaped') curve to describe cumulative adoption.
3. *mixed influence model*; assumes that decisions to adopt are influenced by both societal-external and societal-internal sources. As with the internal influence model, this model of diffusion is represented by a logistic, s-shaped curve.

As with the choice between heterogeneity and homogeneity models, the choice within the communications-based models is a potentially contestable matter. Whilst the external-influence-only model has not been widely used in recent times, several researchers have fit internal-influence-only models to explain technology diffusion data (Teng, Grover & Güttler, 2002; Meade & Islam, 1998). Indeed, Meade & Islam recommend deriving forecasts of technology adoption by combining the outcomes from a range of contrasting models (1998, pp. 1124ff).

This research will use the mixed influence model of diffusion. This is for the following reasons:

- It seems intuitive that the decisions of professional staff in schools to enter candidates for e-assessed qualifications would be influenced by both external and internal communications. School staff receive communications from bodies such as Becta (the British Educational Communications and Technology Agency), awarding bodies, local authorities and so on. Staff also speak to colleagues in their own and neighbouring schools, and it is reasonable to suppose that such communications influence their choice to adopt e-assessment. Using a mixed model of innovation diffusion will permit a researcher to model the relative influence of such external and internal communications.
- The most widely used mixed influence model – the Bass model – is backed up by a wide range of research and management applications (see Mahajan, Muller & Wind's (2000a) edited collection).

The Bass diffusion model

The Bass diffusion model can be defined by the following equation:

$$F(t) = \frac{1 - pe^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}}$$

Equation 1: Cumulative form of the Bass model equation

Where: *F(t)* is the portion (fraction) of the potential market that has adopted up to and including time *t*, *p* is the coefficient of innovation (or 'external influence' – Muller, Peres & Mahajan, 2007) and *q* is the coefficient of imitation (or 'internal influence' – *ibid.*). This version of the formula is that derived by Srinivasan & Mason (1986), and is the optimal version of the model to use for forecasting[iv].

Equation 1 does not refer to a third parameter that is often included in the Bass model formula; that is *m* – market potential (Mahajan, Muller & Wind, 2000b, p. 4). In effect, therefore, all the forecasts in this article pertain to the speed of diffusion of e-assessment and the shape of the diffusion curve, and not to the overall magnitude of diffusion. This is different from model applications that many firms introducing a new product into a commercial market would want; they would be interested in forecasting both how fast high sales would be achieved **and** what the final potential sales would be. This is because it is believed that e-assessment will *eventually* penetrate all schools in England; just as pretty much all schools have *some* computers in them – even if it may have taken a fair amount of time for that state of affairs to come to pass. This would contrast with the situation where a product category was being developed and its producers did not know in advance what proportion of consumers within a total market might adopt it.

The cumulative form of the Bass model function yields an s-shaped curve, as shown in the following figure:

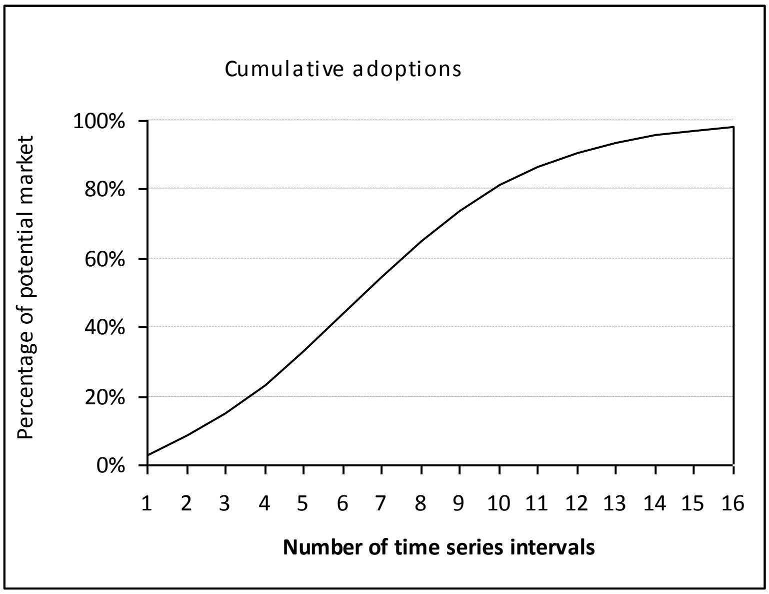

Figure 1: Typical s-shaped curve derived from Bass model

The curve illustrates a function that first ascends slowly (when diffusion is mainly externally influenced), before taking off and commencing a rapid ascent in the number of cumulative adoptions. The central point in a symmetrical Bass model curve is the point of inflection and corresponds to the time interval in which the maximum number of adoptions occurs in the non-cumulative expression of the model (Bass, 1969, p. 219). Rapid ascent in the cumulative number of adopters then continues until a slow-down point is reached (typically a few percentage points less than 100 per cent of all potential adopters). Some researchers have even observed a trough or saddle in adoptions, in which the cumulative number of adoptions actually falls slightly following the slow down period (Hauser, Tellis & Griffin, 2006, p. 691).

By definition, Bass model forecasts relate to a 'product category' rather than a specific brand (see Bass, 1969, p. 215). Thus, this model is suitable to forecast the diffusion of the category 'e-assessment in GCSEs' rather than the sales of a particular awarding body.

p and *q* parameters

There are two issues to be addressed relating to the *p* and *q* parameters. Firstly, the substantive meaning of these parameters and secondly, matters surrounding how to estimate their values for the purposes of forecasting the diffusion of innovations. These two issues are related, but they will be addressed in turn.

In Bass's original exposition of the model he labels *p* and *q* as, respectively: the coefficients of innovation and imitation. He makes it plain that initial purchases of a product (adoptions of an innovation) are influenced by both influences (Bass, 1969, p. 217). External influences are likely to be factors such as mass media and generic communications to the group of potential adopters, whereas internal spurs to adoption were referred to – in early renditions of the Bass model at least – as 'word-of-mouth' factors (Bass, 2004, p. 1835). Some have said that this dual sourcing of influences to adoption is an analogy to the way in which contagious infections spread through a community; carriers of the infection being akin to external influence and those susceptible to catching the disease (although not themselves carriers) being akin to potential adopters who are influenced by the *q* parameter (Bass, 1969, p. 215; Rogers, 2003, p. 335).

Van den Bulte & Stremersch (2004) have theorised social contagion as a concept relevant to the diffusion of innovations. They define social contagion as:

> [the fact] that actors’ adoptions are a function of their exposure to other actors’ knowledge, attitudes, or behaviours concerning the new product. (2004, p.530)

They go on to suggest that:

> Researchers have offered different theoretical accounts of social contagion, including social learning under uncertainty, social-normative pressures, competitive concerns, and performance network effects (*ibid.*)[v]

Muller, Peres & Mahajan (2007) have broadened the definition of the *q* parameter from 'word-of-mouth effects' to include 'social interdependence of all kinds, not only interpersonal communications' (2007, p. 4). As such, *q* includes word-of-mouth communications, but also 'signals' which are 'any market information other than personal recommendation that can be used by a potential adopter to make an adoption decision' (*ibid.* at p. 5). Examples of such signals would be perceptions that products were highly desirable on launch – this perceived popularity being a spur to adopters to imitate their fellow population members and thus adopt the innovation. Muller, Peres & Mahajan's final branch of the broadened conception of internal influence is 'network externalities'; that is, products whose utility increases to an individual adopter as more members of society adopt (*ibid.* at p. 6). Common examples of network externalities include the telephone – not much use if you are the only owner of a phone, more fun if there are two or more of you, or DVDs – the more users of DVD technology there are, the more suppliers are likely to provide a range of movies in that format. Correspondingly, the more movie titles that are available in DVD format, the more an imitative individual's propensity to adopt DVD technology is likely to increase.

Finally, in discussing the substantive meaning of *p* and *q*, it is worth mentioning some aspects of the relationship between them. Firstly, the two parameters are neither logically mutually exclusive nor *necessarily* in a mathematically reciprocal arrangement (i.e. the higher *p*, the lower *q* **must** be, or vice versa). Rather, it can be useful to think of the ratio of *q/p*. This ratio amounts to a 'shape parameter'; possible combinations giving rise to differently-shaped curves, for example:

high *p* – high *q* short time to take off followed by steep rise to diffusion throughout the population

low *p* – high *q* long time to take off followed by rapid increases in adoption once take off had finally been achieved.

The second issue concerned with *p* and *q* parameters is how to estimate them. The first possibility is to use regression techniques on time series data expressing existing adoptions of an innovation. From several alternatives, Putsis & Srinivasan state that non-linear least squares estimation techniques (as described by Srinivasan & Mason, 1986) is now the '*de facto* standard in diffusion research' (2000, p. 285).

However, such techniques depend upon the researcher having a substantial time series of data from which to compute the parameters. Such data are – by definition – not available in the case of a forecast of adoption made at an early stage in an innovation's diffusion process. Various possibilities exist for this situation; for example, Putsis & Srinivasan (2000, p. 281) suggest transforming the Bass model equation into a Bayes format and conditioning *p* and *q* to create a prior distribution. The other possibility is what Bass refers to as 'guessing by analogy' (2004, p. 1835).

In order to support guessing by analogy, it is useful that large databases of empirically-observed *p* and *q* parameters exist. Sultan, Farley & Lehmann's widely cited meta-analysis involved 213 sets of parameters reported in 15 articles (Sultan, Farley & Lehmann, 1990). They found a mean value for the co-efficient of external influence to be .03, whilst the mean for the *q* parameter was .38. Models fitted to data from European countries had higher *q* parameters than those based on US data (*ibid.* at p. 75). The *q* parameter displayed considerable variation amongst reported research exercises (*ibid.*). In their more recent review of diffusion research, Chandrasekaran & Tellis (2008, p. 42) cite a range of different studies and report the parameters derived from them. The following list summarises Chandrasekaran & Tellis' (2008) digest of parameter values – it is worth noting that the different studies suggest slightly different parameter values, especially in the case of *q*:

Coefficient of innovation (*p*)

- The mean value of the coefficient of innovation for a new product lies between 0.0007 and .03.

- The mean value of the coefficient of innovation for a new product is 0.001 for developed countries and 0.0003 for developing countries.
- The coefficient of innovation is higher for European countries than for the United States.

Coefficient of imitation (*q*)

- The mean value of the coefficient of imitation for a new product lies between 0.38 and 0.53.
- Industrial/medical innovations have a higher coefficient of imitation than consumer durables and other innovations.
- The mean value of the coefficient of imitation for a new product is 0.51 for developed countries and 0.56 for developing countries.

Van den Bulte & Stremersch cite findings of a meta-analysis of 746 sets of parameter estimates reported in 54 publication (2004, p. 536). They evaluate the ratio *q/p* in terms of hypothesised aspects of social contagion: social learning under uncertainty, social-normative pressures, competitive concerns, and performance network effects (2004, p. 540). They find, amongst other things that 'more collectivist cultures have a higher *q/p* ratio' (*ibid.*). They account for this finding by suggesting that cultures that emphasise conformity to social norms and group behaviour will have a higher *q*-to-*p* ratio. Also, they suggest that cultures with high 'power distance'[vi] will find internal influence being higher in relation to external influence. They also report the – to them – surprising finding that uncertainty-avoiding cultures have a *lower q/p* ratio. Despite this finding, however, they also find – more consistently with the overall trend of their findings – that uncertainty-avoiding cultures return to a higher *q/p* ratio when presented with an innovation that has competing standards.

Teng, Grover & Güttler (2002) reviewed diffusion patterns reported for 20 ICTs. Their first finding that is relevant for present purposes was that they suggested that adoption of ICTs was largely a process of imitativeness rather than innovation, and – correspondingly, *q* could be expected to be relatively high, but *p* would be low (between 0.0001 and 0.0062) (2002, p. 19). They also clustered the ICTs according to their *p* and *q* values, and then posited substantive explanations for the innovativeness and imitative characteristics that were relevant to each cluster.

Teng, Grover & Güttler's substantive contentions that are relevant for current purposes include:

- information technologies … that depend on market externality[vii] will diffuse more slowly than those that do not have this constraint
- tool technologies that can be directly utilized as a support tool by users will have lower set-up time and, hence, diffuse faster than systems technologies that require extensive professional development or training in adoption and implementation of the technology (2002, p. 24)

Possible objections to the Bass model

There are several lines of reasoning that could be advanced to argue against using the Bass model to derive forecasts for the diffusion of e-assessment. These are considered in this section.

Firstly, one might object to the use of a marketing model at all to forecast the diffusion of e-assessment. It may be felt that it is too far a leap from high-definition TVs, air conditioners, clothes dryers and so on to e-assessment used in GCSE exams. To counter this objection, it may be observed that – although Bass modelling did indeed originate in predicting the adoption of purchase of domestic 'white' appliances – it has been used much more widely throughout its 40-year development. For example, it has been widely used to forecast diffusion of medical innovations such as CAT scanners or mammography technology (see: Sultan, Farley & Lehmann, 1990, p. 72). A model similar to Bass's has even been used to predict the diffusion of a range of educational innovations such as: 'modern mathematics', foreign language teaching in elementary schools, accelerated programs, language laboratories and team teaching (Lawton & Lawton, 1976). Given this breadth of use of aggregate diffusion models, the extension to the case of e-assessment seems reasonable. Also, although the adoption of e-assessment is not primarily influenced by cost factors, it *is* a purchase decision and thus more suitable to first purchase aggregate diffusion model analysis than the adoption of team teaching, modern mathematics and so forth.

The second issue concerns the simplicity vs. the representativeness of the Bass model. To its supporters the Bass model is an elegant and parsimonious empirical generalisation (Bass, 1995). Others (whether outright critics of the Bass model, or those who have extended and generalised it) have made various criticisms. Such criticisms include:

- incomplete reporting of necessary features of model use – for instance non-reporting of error terms associated with regressions to the model equation (Parker, 1994, p. 360)
- uses of the model in which assumptions of the model were violated significantly enough to make its application invalid (Meade & Islam, 2006, p. 521), and
- lack of flexibility in the core model that – it is suggested – need to be rectified by adding more parameters into the equation (Mahajan, Muller & Bass, 1990, pp. 11ff.).

The current research's response to this second set of possible criticisms of the Bass model is as follows: firstly, it is believed that all relevant matters that potentially challenge the validity of the application of the model to the current circumstances have been made explicit. It is the view of this researcher that any challenges to the validity of application are not sufficient to disbar that application. Indeed, it is further believed that sufficient information has been made available so that readers can make their own minds up about the validity of model application. Secondly, it is acknowledged that this article proposes the application of the core form of the Bass model – minus any of the bells and whistles that could potentially be used to model some of the complexities of the reality of e-assessment diffusion. This is done mainly because this is the first application of a Bass model to e-assessment diffusion in the literature. As such, there is a parsimony and associated clarity that can be gained by using the core model. If this research is thought to have a sound basis, then subsequent projects may come along to fine tune the use of the forecasting tools. In the 'further work' section of this paper some potential extensions and generalisations of the core model are evaluated.

The third set of concerns relates to the power of the Bass model (and indeed any forecasting technique) when based on very limited data. Several reviews have pointed out that – for the most valid application – the Bass model needs to be estimated from data that contain both the take off and slow down in sales (adoptions) (see: Mahajan, Muller & Bass, 1990, p. 9; Chandrasekaran & Tellis, 2008, p. 49). Waiting for such data to become available would increase the technical validity of the model, but diminish its utility as a forecasting tool almost to zero.

A specific issue with the Bass model is an observation of 'left-truncated' adoption data (see: Jiang, Bass & Bass, 2006). Left-truncated data can be illustrated in Figure 2 of non-cumulative adoption of an innovation.

If the diffusion of an adoption at period 4 in Figure 2 (see vertical line on figure) was interpreted as though time period 4 were the first adopting period, then Bass model parameters would potentially be inaccurately estimated. This state of affairs could occur when a new forecast was relying on data describing previous innovations to estimate parameters for the new forecast. In such cases, the forecast could underestimate the amount of time that the new innovation was likely to take to diffuse, since it would not account for 'shadow' diffusion[viii] (Rogers, 2003, p. 354) of the earlier innovation.

Jiang, Bass & Bass (2006) assume that left-truncated adoption data would relate to earlier innovations that were standing as analogies for the new innovation (whose diffusion was to be forecast), and that the starting point for the new innovation would be clear. In the current case, the potential problem is the reverse one; it is intended to

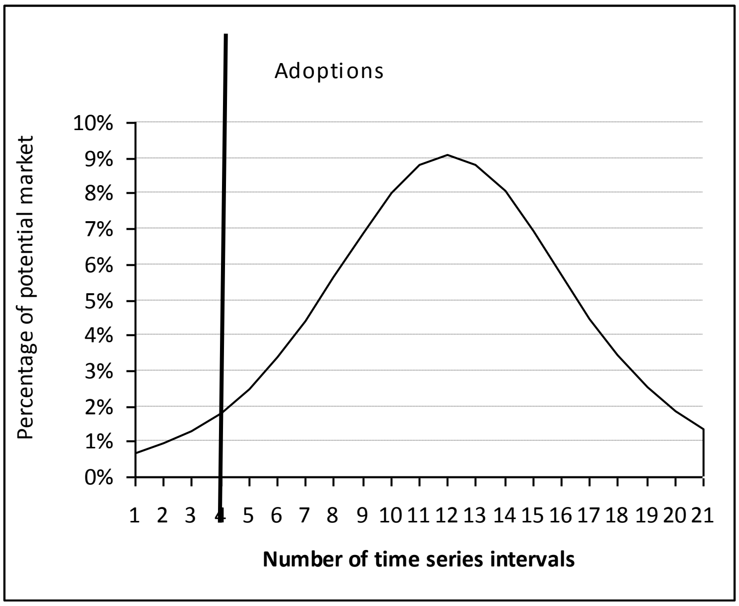

Figure 2: Non-cumulative form of Bass model showing left-truncated data

model diffusion of e-assessment in GCSEs. However, it is known that e-assessment has been used for various kinds of tests and exams in England before the diffusion of the assessment method in GCSE. If prior use amounts to left-truncation of the new non-cumulative diffusion curve, then the period taken to diffuse the innovation may be overestimated.

The potential flakiness of Bass model predictions in the presence of little or no data is acknowledged. However, it is believed – following Parker, 1994, p. 367 – that using an explicit, replicable methodology to construct forecasts at least permits the issue of e-assessment diffusion to be debated in a structured context. The unattractive alternative is for e-assessment enthusiasts and sceptics to assert, respectively, the imminent ubiquity or long-term shallow penetration of e-assessment, and for those assertions to be unfalsifiable and for thinking in this area consequently not to advance. Whilst forecasts may not – ultimately – be right, the collection of data can also be planned and the more data are collected, the more accurate forecasts can become.

Thus, it is submitted that, although at least three substantive sets of objections to the use of the Bass model can be made, these objections can be overcome. Therefore, this research uses the Bass model to make forecasts about the future diffusion of e-assessment in GCSEs. The methods used to do this are described in the next section.

Method

This research uses the cumulative version of the core Bass model equation set out in equation 1 above to posit forecasts of the diffusion of e-assessment in GCSE qualifications. It reports each forecast in the form of an s-shaped logistic curve and a table showing the percentage of potential adopters forecast to adopt over time. Also, the results section will contain comment upon features of the diffusion curve.

In order to give a comprehensive picture of various possibilities that can be forecast for the diffusion of e-assessment, three forecasts are made.

1. A forecast using the values of *p* and *q* found by Sultan, Farley & Lehmann (1990) in their large-scale review (0.03 and 0.38, respectively). This forecast is considered as a 'control condition'; given that the Sultan, Farley & Lehmann's values for *p* and *q* can be considered as typical of large numbers of innovations.
2. A forecast which is based on *p* and *q* parameters that are derived by using non-linear least squares regression to relate data from previous e-assessment diffusions to the cumulative Bass model equation.
3. A forecast that considers the substantive nature of e-assessment as an innovation, and staff in schools as a potentially adopting population. Using this substantive argument (being akin to what the literature refers to as 'management judgement'), this forecast posits 'intuitive' values for *p* and *q*.

Once the three forecasts derived, they are discussed. This discussion highlights substantial implications of the findings as well as simple timings. Finally, the article reviews some of the potential extensions and generalisations to Bass modelling and suggests those which might profitably be pursued in further research.

Data

The data required for each forecast are very different. They are summarised below.

1. The first forecast implements the mean values of *p* and *q* found by Sultan, Farley & Lehmann (1990). Thus, the only data needed are the two parameter values. Of course, Sultan, Farley & Lehmann (1990) themselves processed large quantities of data in their review.

2. The second forecast is based on empirical data of e-assessment diffusion in another country. It uses the following series of adoption data that describe the diffusion of e-assessment in senior secondary education/pre-university education national final examinations in the Netherlands[ix] (based on a total of 500 schools that could potentially use e-assessment for the exams).

Year	**Number of schools adopting**	**Number of schools as proportion of total number of schools**
2003	10	0.02
2004	50	0.1
2005	100	0.2
2006	300	0.6
2007	380	0.76

Table 1: Diffusion of e-assessment in Dutch examinations

3. The third prediction takes known substantive features of e-assessment diffusion (for example from previous trials of e-assessment for high-stakes tests in England) and analyses these in the light of substantive features that are described in the literature as influencing e-assessment adoption.

Results

Sultan, Farley & Lehmann (1990) mean values for **p** *and* **q**

The mean values for the two parameters found by Sultan, Farley & Lehmann (1990) were, respectively, 0.03 and 0.38. When applied to the cumulative Bass model as defined in this article, with a starting year of 2008, the following diffusion pattern is forecast.

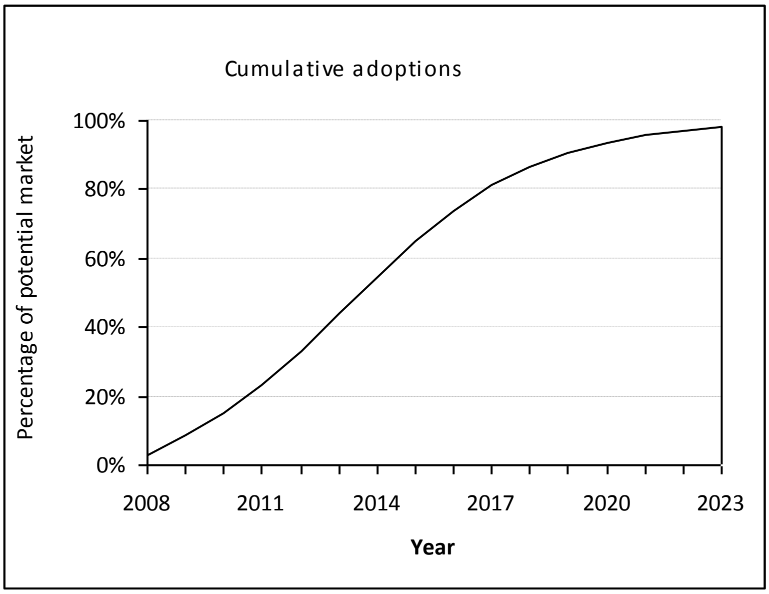

Figure 3: Forecast e-assessment diffusion using Sultan, Farley & Lehmann mean values

Year	Cumulative adoptions (as percentage of total market)
2008	3.00
2009	8.51
2010	15.05
2011	23.32
2012	33.12
2013	43.92
2014	54.90
2015	65.17
2016	74.07
2017	81.28
2018	86.81
2019	90.87
2020	93.76
2021	95.78
2022	97.16
2023	98.10

Table 2: Forecast e-assessment diffusion using Sultan, Farley & Lehmann mean values

Features of this diffusion pattern include the following:

- By 2015, the year that the archetypal child who inspired this article would take her GCSEs, approximately 65 per cent of schools would be offering e-assessed GCSEs in at least one subject.
- The peak of diffusion in the non-cumulative Bass model, and the point of inflection in the cumulative version of the model will be shared between 2013 and 2014.
- By 2020, approximately 94 per cent of all centres will be running GCSEs on-screen.

Model using parameters estimated from Dutch data

A colleague from the Dutch national institute for educational measurement kindly provided an estimate of the number of exams centres that had adopted e-assessment over the past few years. Those data are summarised in Table 1. Non-linear least squares regression was applied to those data to estimate values for the *p* and *q* parameters. The values of those parameters, and associated confidence bands, are shown in Table 3.

			95% Confidence Interval	
Parameter	Estimate	Std. error	Lower bound	Upper bound
p	.010	.006	-.009	.030
q	1.229	.197	.601	1.857

Table 3: *p* and *q* parameter values derived from data from the Netherlands

These parameter estimates were added to the Bass model to give the curve in Figure 4 and data table 4.

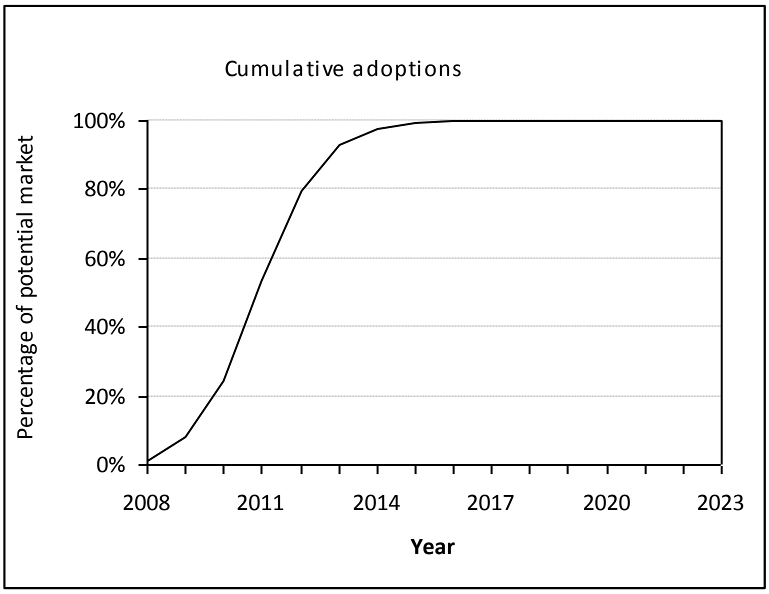

Figure 4: Diffusion modelled using data from the Netherlands

Year	Cumulative adoptions (as percentage of total market)
2008	1.00
2009	8.10
2010	24.47
2011	53.23
2012	79.79
2013	93.18
2014	97.92
2015	99.39
2016	99.82
2017	99.95
2018	99.99
2019	100.00
2020	100.00
2021	100.00
2022	100.00
2023	100.00

Table 4: Diffusion modelled using data from the Netherlands

- By 2015, cumulative adoptions will have reached over 99 per cent of all GCSE centres.
- The peak of diffusion in the non-cumulative Bass model, and the point of inflection in the cumulative version of the model will in 2011.
- Between 2015 and 2020, under this parameterised diffusion model, the remaining fraction of one percent of exams centres that had not already adopted e-assessment for GCSEs would do so.

These results – derived by analogy to the diffusion of e-assessment in the Netherlands – suggest a very rapid diffusion. Indeed, the *q* parameter derived from Dutch data looks quite high; somewhat outside the most commonly found values in Chandrasekaran & Tellis' 2008 review (summarised above). Also, the 95 per cent confidence interviews for the *q* parameter are quite high and it may be that interpreting *q* as falling towards the lower confidence bound (just above .6) might give a more intuitive result. However, the Dutch data are direct data of e-assessment adoption, and thus the *q* estimate returned by the regression process is retained despite some reservations that this might give an overly 'optimistic' forecast.

Substantive argument relating to the population of potential adopters

This forecast will be based on *p* and *q* parameters that are derived by a substantive argument as to the nature of the population into which e-assessment seeks to diffuse; that is, English schools. Those features of English schools will be linked to factors likely to affect the values of *p* and *q*, which were set out in the literature review section above.

Firstly, one might consider whether network externalities are likely to apply to e-assessment. It does seem likely that e-assessment diffusion could be subject to network externalities in the same way that DVD diffusion is. That is, the more examination subjects and syllabuses on offer, the more likely an imitative person (or organisation) would be to adopt. If this scenario is accepted as accurate, then it means that the ratio of *q/p* will be relatively high.

Secondly, one might consider the extent of hierarchy amongst potential adopters of e-assessment in English schools. Clearly, staff in schools are less conscious of formal hierarchies than some groups (e.g. uniformed services such as the military, police and fire brigade). However, it is reasonable to assert that the consciousness of hierarchies would affect the potentially adopting population when they make their choices with respect to the innovation e-assessment. There are formal hierarchies amongst the teachers who contribute to that decision – for instance between school senior leaders, heads of department and main-scale teachers. Also, there is the element of hierarchy that often exists between non-teaching staff (e.g. exams officers, network managers and ICT technicians) and teachers. As such, it is reasonable to suppose that e-assessment adoption in the case of GCSEs is a scenario in which a high power distance plays a role. It further follows from this that a high *q/p* ratio would pertain.

Thirdly, those concerned with education in England, and specific institutions such as government, regulators, ABs and schools are risk averse when dealing with high-stakes examinations and qualifications (Richardson, 2007). Also, GCSEs are – unusually in the international context – provided by five awarding bodies. Work is underway to try to provide a standard platform upon which e-assessments from different ABs can run (QCA, 2007). However, it seems likely at the time of writing that at least two delivery platforms for e-assessment will exist in England in the foreseeable future. Therefore, these two scenarios suggest a combination of risk aversion and a technology that has competing standards. As with previous elements of this forecast-by-substantive-argument this suggests a high *q/p* ratio.

Finally, in this part of the argument, one can consider whether e-assessment is a support tool or a system technology. The best response to this question seems to be that e-assessment contains elements of both a support tool for users and systems technology. The actual software with which an exam candidate sits the test amounts to a tool technology. However, any credible software solution for e-assessment is likely to include a substantial 'back-office-processing' module as well (Boyle, 2005, p. 36). Such a module would perform tasks such as: registering candidates for tests, allocating candidates to test sessions, collecting back test data (responses to questions) and sending them on to an awarding body central server. It would take substantial amounts of time and expertise to set up and maintain (*ibid.*). As such, although e-assessment is a tool technology, it is also a major systems technology and as such is likely to diffuse relatively slowly.

From the four strands of argument advanced above it is clear that the ratio of *q* over *p* will be high. A high *q/p* ratio could mean:

- High *q*/average *p*
- Average *q*/low *p*
- High *q/* low *p*

It is proposed that the drift of these arguments about the nature of the population of potential e-assessment adopters has been to stress that population's relative conservatism and lack of innovativeness. Therefore, it is proposed to select a combination with an average value for *q* and a low value for *p*. This can be achieved by taking values suggested by Chandrasekaran & Tellis (2008, p. 42). In Chandrasekaran & Tellis' summary, the mean value for *q* for a new product was 0.51 for developed countries. The lowest value found for *p* was 0.001. These values have been input to create the chart in Figure 5 and data table 5, below:

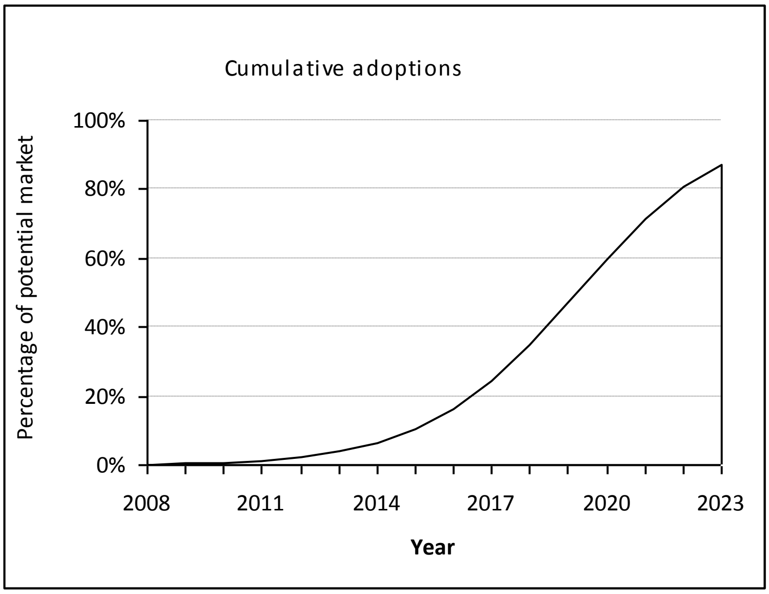

Figure 5: e-assessment diffusion resulting from substantive argument

Year	Cumulative adoptions (as percentage of total market)
2008	0.10
2009	0.35
2010	0.71
2011	1.30
2012	2.27
2013	3.85
2014	6.37
2015	10.29
2016	16.14
2017	24.37
2018	35.00
2019	47.34
2020	60.00
2021	71.44
2022	80.66
2023	87.43

Table 5: e-assessment diffusion resulting from substantive argument

Features of this diffusion pattern include the following:

- By 2015, cumulative adoptions will have reached only 10 per cent of GCSE centres.
- The peak of diffusion in the non-cumulative Bass model, and the point of inflection in the cumulative version of the model will be in 2020.
- By that time (2020) only approximately 60 per cent of GCSE centres will be using e-assessment.

Discussion

Comparison of the three forecasts

The extent to which the three forecasts for e-assessment diffusion in GCSEs diverge can be illustrated from the curves in Figure 6.

The three forecasts diverge substantially, and thus do not provide a 'single right answer'. Rather, they provide a range of forecasts; from the most ‘bullish’ prediction of e-assessment diffusion for GCSEs to the more cautious. The assumptions underlying each forecast are set out explicitly in this article. Table 6 summarises the strengths and weaknesses of the approach underlying all the forecasts, so as to aid their evaluation.

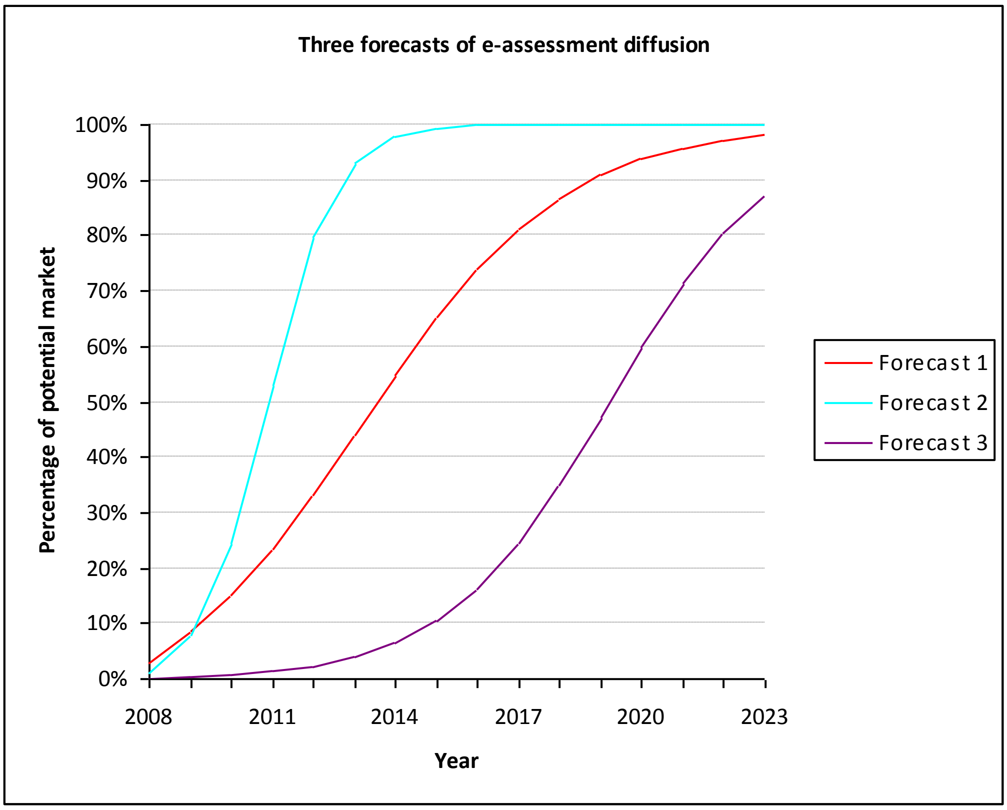

Figure 6: The three forecasts of e-assessment diffusion on a single figure

Forecast	Strengths	Weaknesses
Sultan, Farley & Lehmann meta-analysis	• parameters derived from broad meta-analysis • Sultan, Farley & Lehmann analysis highly regarded within Bass model tradition	• These parameters have no link to education.
Dutch e-assessment data	• based on actual e-assessment diffusion data	• *q* parameter seems on the high side of the plausible range • *q* has large error term
Substantive argument	• based on a substantive argument that adapts diffusion research to the specific circumstances of e-assessment use in GCSE	• not derived from empirical diffusion data • Contrary arguments could be validly made.

Table 6: Strengths and weaknesses of the three forecasts

Whilst this article has been written, ABs have agreed to provide a regular data feed to regulators so that a time series can be built up showing the diffusion of e-assessment in GCSEs, which is helpful, since diffusion forecasts increase in accuracy as such time series are built up (Parker, 1994, p. 368). As such cumulative data are built up, the model above which best describes e-assessment diffusion should become progressively clearer.

Possible further research

In this final section of the paper, some possible directions for further research are suggested. These possible directions all involve the application of Bass model generalisations or extensions. Such add-ons to the core model are not proposed merely as 'bells and whistles', but rather as ways of addressing key substantive issues that affect the diffusion of e-assessment.

The first issue is the 'fuzzy left-hand side' of the e-assessment diffusion curve. Although there has been very limited use of e-assessment in GCSEs to date, there has been use of e-assessment in other contexts. These include: high-stakes tests in other countries (especially the USA), low-stakes tests (often commercial products that can be bought by schools) as well as other formal testing in the United Kingdom such as the theory test for drivers and riders or the Qualified Teacher Status ICT, numeracy and literacy skills tests. It would be reasonable to argue that such prior uses of e-assessment ought to have some effect on its use for GCSEs and thus be modelled in an extension or generalisation of the core Bass model. There are several ways in which this could be done.

Firstly, one or more of the applications of e-assessment described above could be considered as a prior generation of a technology and thus contribute to an analysis of e-assessment diffusion in GCSEs using the successive generations diffusion model proposed by Norton & Bass (1987). Alternatively, one might consider that the lack of clarity as to when e-assessment as an innovation actually started out might lead to the potential issue of left-truncation of data. As such, the Virtual Bass Model (VBM), proposed by Jiang, Bass & Bass (2006) might be fruitfully applied.

However, whilst both those approaches might use more available data and thus provide enlightening findings, an underlying assumption of these extensions to the core Bass model may limit their usefulness. Essentially, both the VBM and the successive generations model will seek to incorporate additional information into an expanded model in order to produce a better forecast of diffusion. However, neither will investigate the substance of the processes that take place between the introduction of an innovation and its take off. This may be an omission; after all, it is possible that e-assessment use could become 'entrenched' in a small number of small cohort GCSEs, but not be taken up across many large cohort subjects. Golder & Tellis (1997) have provided a model of the period of the innovation diffusion curve between introduction and take off which may be useful to investigate this phenomenon.

A further line of diffusion modelling scholarship that might profitably be employed would be that which considers the effect of supply constraints on the diffusion of an innovation. There are several constraints on the use of e-assessment in GCSEs. For example, regulatory arrangements may require ABs to get approval for an amendment to the assessment method before its introduction. Also, regulators may require ABs to demonstrate the measurement properties of an e-assessment (e.g. the comparability of score results between parallel pencil-and-paper and e-assessed versions – see: Raikes & Harding, 2003); this may slow down the introduction of the innovation. Equally, the ABs themselves may be slow to introduce e-assessment into GCSEs. This may be for a range of reasons, such as: the bodies' intrinsic risk aversion or the fact that resource is deployed elsewhere – for instance in the development of new style qualifications.

The research of Mesak & Darrat (2003), Ho, Savin & Terwiesch (2002) and Kumar & Swaminathan (2003) (amongst others) provides a framework within which to investigate the diffusion of e-assessment under supply constraints.

Whilst it is important to model how the need to demonstrate reliability, validity and comparability may operate as constraints on the diffusion of e-assessment in GCSEs, it is also important to consider the effect that some diffusion patterns could have on the ability to demonstrate the necessary measurement properties of e-assessment. In particular, a diffusion pattern with a long period between introduction and take off, followed by a rapid ascent from take off to slow down (a common pattern for technology diffusion – Teng, Grover & Güttler, 2002, and possibly exemplified by the 'substantive argument' forecast in the current research) could be particularly problematic for high-stakes e-assessment. This would be because, during the initial slow diffusion period before take off, it may be difficult for ABs to find participants willing to take part in studies to establish validity, reliability and comparability. However, once the take-off point has been passed, suddenly there would become very many adopters of e-assessed GCSEs. However, if it had not been possible to establish validity, reliability and comparability in the pre-take-off period, then this large new group of candidates risk sitting exams whose measurement properties have not been established.

The final extension to the classic Bass model that could provide light on e-assessment adoption would be an approach to model the use of e-assessment by multiple candidates in a single school. As discussed above, the origins of the Bass model were in modelling the diffusion of infrequently purchased, 'white goods'. As such, the number of sales of a product were taken as equal to the number of adoptions. However, many GCSEs are intended to be taken by 'whole cohorts' of candidates (an entire school year group for example). Research into a pilot of an e-assessed national curriculum test (Quinlan & Boyle, 2005) suggested that even though the e-test was intended to be taken by whole cohorts of pupils, in fact, in many schools, much lower proportions of children took the e-test. In Quinlan & Boyle's 2005 research, from 402 schools for which data were available, 224 schools entered 60 per cent or more of their pupils for the e-test. However, 116 schools entered less than 40 per cent of their full cohort for the test. Reasons for schools entering less than their full cohort of pupils included: schools lacking sufficient workstations to enter all pupils, schools – although having sufficient workstations – not wishing to 'tie up' all their IT facilities on a test, and schools' IT networks not being sufficiently powerful to transmit many large packets of test data simultaneously (Boyle, 2006).

Given this need to assess the likelihood of schools testing whole cohorts in e-assessed GCSEs, it is useful that literature exists to forecast repeat sales of products (e.g. Fader & Hardie, 2001). Indeed, Teng, Grover & Güttler's observation (2002, p. 23) of a high correlation between one use of an IT innovation within an organisation (diffusion), and multiple uses of that product across the organisation (infusion) is relevant. Also important is their supposition (*ibid.*) that infusion may be describable using a similar s-shaped curve to those that describe diffusion. However, it may be that whole cohort e-assessment for GCSEs presents a special challenge to this research strand – in that perceptions of total infusion not being possible could feed back to constrain diffusion (for instance if schools said 'if we can't test *all* the kids on screen, then we won't test *any* of them').

Conclusion

This article argued that e-assessment diffusion was a matter worthy of substantive study. It showed how differences in diffusion might affect interested parties (students, awarding bodies, the regulator). It also supposed that one could understand diffusion by thinking about the 'demand side'; specifically about how quickly schools and other examination centres would be likely to adopt e-assessment for GCSEs.

To address its core issue, the article cited research literature to propose the use of a well-known diffusion model. That model assumed a homogeneous population of potential adopters and that – whilst external communications did explain a certain amount of adoption – imitative aspects of communication (broadly defined) can explain a larger part of adoption.

The model was parameterised to give three diffusion curves. The three derived curves diverged substantially, and it is not possible – by the end of this article – to say which forecast is the most likely to come true. However, strengths and weaknesses of the forecasts have been made explicit, to aid evaluation. Further data collection is planned and this should allow researchers to become clearer over time as to the most likely diffusion curve for e-assessment in GCSEs.

At the end of this article, it remains the view of the researcher that it is worthwhile to investigate e-assessment diffusion by thinking about how schools take up the innovation. It can be too easy to simply think in terms of policy initiatives, marketing drives and so forth – forgetting that schools always have a choice about whether or not to implement e-assessment for their students.

It is further argued that the research reported here shows that it is feasible and informative to use a well-known diffusion model. Further, it is suggested that it is good practice to use the core version of the model for the first instance of research into e-assessment diffusion. Doing so allows clarity as to the meaning of findings, without the presence of too many confounding variables. Once this research has been accepted, replicated or forecasts have been updated as real adoption data come in, it will be possible to add extensions and generalisations to the core model. But this can be done from a position of understanding of the functioning of the basic model. Further, the use of a demand-side, homogenous-population diffusion model in this instance does not preclude the subsequent use of entirely different models – perhaps derived from economics literature.

This research may be a valid response to Oates' (2007) wish that assessment research be informed by broader social science. Further, whilst diffusion modelling has been illustrated in the context of e-assessment, it may also have a use in other areas relevant to qualifications and exams, such as the adoption of new qualifications, reasonable adjustments and so on.

About the Author

Andrew Boyle is the Head of Assessment Research at the Office of the Qualifications and Examinations Regulator (Ofqual) in England. He has been researching e-assessment for about 12 years, having an interest in areas such as: Computerised Adaptive Testing (CAT), the marking of natural language by machines, the use of sophisticated items in e-assessment, quality assurance documents in e-assessment and the formative use of e-assessment.

Acknowledgements
Thanks are due to the following Ofqual staff: Carmel Fung for great assistance with the non-linear regression, Paul Newton, Tina Isaacs and Dennis Opposs for content reviews. Thanks are also due to Peter Hermans of CITO-Groep, the Netherlands for providing the data regarding the numbers of Dutch schools using e-assessment. Portia Isaacson Bass provided helpful advice relating to the Bass model.

Sources:
Bass, F.M. 1969. "A new product growth model for consumer durables". *Management Science,* 15, 215-227.

Bass, F.M. 1995. "Empirical generalizations and marketing science: a personal view". *Marketing Science,* 14(3), part 2 of 2, G6-G19.

Bass, F.M. 2004. "Comments on 'A new product growth for model consumer durables'". *Management Science*, 50(12) Supplement, December 2004, 1833-1840.

Boston, K. 2004. *Speech to e-assessment conference*. Tuesday, 20 April 2004, Royal Festival Hall. Available online at: http://www.qcda.gov.uk/5465.aspx.

Boyle, A. 2005. *Interim evaluation of the 2005 pilot of the Key Stage 3 ICT tests*: a report to the Department for Education and Skills. Available online at: http://www.ofqual.gov.uk/files/qca-06-2340-interim_eval_report2005.pdf.

Boyle, A. 2006. *Evaluation of the 2006 pilot of the key stage 3 ICT test*. Available online at: http://testsandexams.qcda.gov.uk/libraryAssets/media/2006_pilot_evaluation_final_version.pdf.

Boyle, A. 2008. *The regulation of examinations and qualifications: an international study*. Available online at: http://www.ofqual.gov.uk/1726.aspx.

Chandrasekaran, D. and G.J. Tellis. 2008. *A critical review of marketing research on diffusion of new products*. Marshall research paper series working paper MKT 01-08. Available online at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=969775.

Chatterjee, R. and J. Eliasberg. 1990. "The innovation diffusion process in a heterogeneous population: a micromodeling approach". *Management Science*, Sep90, 36(9), 1057-1079.

Department for Children, Schools and Families (DCSF). 2007. *Diplomas*. Available online at: http://www.dcsf.gov.uk/14-19/index.cfm?sid=3&pid=224&ctype=None&ptype=Contents.

Europe Economics. 2008. *Scoping exercise for a study of the efficiency of the qualifications system: a report for QCA*. Available online at: http://www.ofqual.gov.uk/files/scoping_study_october_29_2007.pdf.

Fader, P.H. and B.G.S. Hardie. 2001. "Forecasting repeat sales at CDNOW: a case study". *Interfaces*, 31(3), S94-S107.

Geroski, P.A. 2000. "Models of technology diffusion". *Research Policy*, 29, 603-625.

Golder, P.N. and G.J. Tellis. 1997. "Will it ever fly? Modeling the takeoff of really new consumer durables". *Marketing Science*, 16(3), 256.

Hauser, J., G.J. Tellis, and A. Griffin. 2006. "Research on innovation: a review and agenda for Marketing Science". *Marketing Science*, 25(6), 687-717.

Ho, T-H., S. Savin, and C. Terwiesch. 2002. "Managing demand and sales dynamics in new product diffusion under supply constraint". *Management Science*, 48(2), 187-206.

Jiang, Z., F.M. Bass and P.I. Bass. 2006. "Virtual Bass model and the left-hand data-truncation bias in diffusion of innovation studies". *International Journal of Research in Marketing*, 23, 93-106.

Joint Information Systems Committee (JISC). 2006. *e-Assessment Glossary* (Extended). Available online at: http://www.jisc.ac.uk/uploaded_documents/eAssess-Glossary-Extended-v1-01.pdf.

Kingdon, M. 2005. *The development of e-assessment 2004 to 2014*: the second of a series of occasional papers on e-assessment issues sponsored by Exam on Demand Ltd September 2005. Available online at: http://www.examondemand.co.uk/Papers/EXoD%20Report%20final%20no2.pdf.

Kumar, S. and J.M. Swaminathan. 2003. "Diffusion of innovations under supply constraint". *Operations Research*, 51(6), 866-879.

Lawton, S.B. and W.H. Lawton. 1976. *An autocatalytic model for the diffusion of educational innovations*. ERIC ID: EA007 875. Available online at: http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/32/c2/34.pdf.

Le Métais, J. 2002. *International developments in upper secondary education: Context, provision and issues*: INCA Thematic Study No. 8. Available online at: http://www.inca.org.uk/pdf/cav_final_report.pdf.

Mahajan, V., E. Muller and F.M. Bass. 1990. "New product diffusion models in marketing: a review and directions for research". *Journal of Marketing*, 54, 1-26.

Mahajan, V., Muller, E. and Y. Wind. 2000. New product diffusion models: from theory to practice. In Mahajan, V., Muller, E. and Y. Wind. (eds.) *New product diffusion models*. Boston: Kluwer Academic.

Mahajan, V., Muller, E. and Y. Wind. (eds.). 2000. *New product diffusion models*. Boston: Kluwer Academic.

Mahajan, V. and R.A. Peterson. 1985. *Models for innovation diffusion: quantitative applications in the social sciences*. Newbury Park, CA: SAGE.

Meade, N. and T. Islam. 2006. "Modeling and forecasting the diffusion of innovation – a 25 year review". *International Journal of Forecasting*, 22(3) 519-545.

Meade, N. and T. Islam. 1998. "Technological forecasting – model selection, model stability, and combining models". *Management Science*, 44, 1115-1130.

Mesak, H.I. 2003. "An empirical inquiry into new subscriber services under independent adoption processes". *Journal of Service Research*, 6(2), 180-192.

Muller, E., R. Peres and V. Mahajan. 2007. *Innovation diffusion and new product growth: beyond interpersonal communications*. Working paper. Available online at: http://www.hitechmarkets.net/files/ReviewPaperAugust2007final.pdf.

Norton, J.A. and F.M. Bass. 1987. "A diffusion theory model of adoption and substitution for successive generations of high-technology products". *Management Science*, 33(9), September 1987, 1069-1086.

Oates, T. 2007. *The constraints on delivering public goods – a response to Randy Bennett's 'What does it mean to be a non-profit educational measurement organization in the 21st Century?'* A paper presented at the International Association for Educational Assessment Annual Conference, Baku, Azerbaijan, September 2007.

Parker, P.M. 1994. "Aggregate diffusion forecasting models in marketing: a critical review". *International Journal of Forecasting*, 10(2), 353-380.

Price Waterhouse Coopers (PWC). 2005. *The market for qualifications in the UK*: final report to: Qualifications and Curriculum Authority, Qualifications, Curriculum and Assessment Authority for Wales, Council for the Curriculum, Examinations and Assessment for Northern Ireland. Available online at: http://www.ofqual.gov.uk/files/pdf_05_2239full_report.pdf.

Putsis, W.P. and V. Srinivasan. 2000. Estimation techniques for macro diffusion models. In Mahajan, V., Muller, E. and Y. Wind. (eds.) *New product diffusion models*, 263-291. Boston: Kluwer Academic.

Qualifications and Curriculum Authority (QCA). 2007. *Fifty eighth meeting of the Board of the Qualifications and Curriculum Authority*: Minutes of the meeting held on Wednesday, 12 September 2007 at Ansty Hall, Nr Coventry, Warwickshire. Available online at: http://www.qcda.gov.uk/libraryAssets/media/Minutes-12-September-2007.pdf.

Qualifications and Curriculum Authority (QCA), Department for Children, Education, Lifelong Learning and Skills (DCELLS), Council for the Curriculum, Examinations and Assessments (CCEA) & Scottish Qualifications Authority (SQA). 2007. *Regulatory principles for e-assessment.* Available online at: http://www.ofqual.gov.uk/files/Final_regulatory_principles_document_-_PRINTED.pdf.

Qualifications and Curriculum Authority (QCA), Qualifications and Curriculum Group, Department for Education, Lifelong Learning and Skills (DELLS) & Council for the Curriculum, Examinations and Assessment (CCEA). 2008. *Annual qualifications market report*. Available online at: http://www.ofqual.gov.uk/files/Annual_qualifications_market_report_28-03-08.pdf.

Quinlan, M. and A. Boyle. 2005. *Comparative analyses of results from KS3 ICT test and other NC tests*. QCA internal report.

Raikes, N. and R. Harding. 2003. "The Horseless Carriage Stage: replacing conventional measures". *Assessment in Education: Principles, Policy & Practice*, 10(3), November 2003, 267-277.

Rangaswamy, A. and S. Gupta. 2000. Innovation adoption and diffusion in the digital environment: some research opportunities. In Mahajan, V., Muller, E. and Y. Wind. (eds.) *New product diffusion models*, 75-96. Boston: Kluwer Academic.

Richardson, W. 2007. "Public policy failure and fiasco in education: perspectives on the British examinations crises of 2000-2002 and other episodes since 1975". *Oxford Review of Education*, 33(2), May 2007, 143-160.

Ripley, M. 2006. *The four changing faces of e-assessment 2006-2016*. Insight observatory for new technologies and education thematic dossiers. Available online at: http://insight.eun.org/ww/en/pub/insight/thematic_dossiers/articles/e_assessment/eassessment2.htm.

Rogers, E.M. 2003. *Diffusion of innovations: fifth edition*. New York: Free Press.

Srinivasan, V. and C.H. Mason. 1986. "Nonlinear least squares estimation of new product diffusion models". *Marketing Science*, Spring 1986, 5(2), p. 169-178.

Sultan, F., J.U. Farley and D.R. Lehmann. 1990. "A Meta-analysis of diffusion models". *Journal of Marketing Research*, 27, 70–77.

Teng, J.T.C., V. Grover and W. Güttler. 2002. "Information technology innovations: general diffusion patterns and its relationships to innovation characteristics". *IEEE Transactions on Engineering Management*, 49(1), February 2002.

Van den Bulte, C. and S. Stremersch. 2004. "Social contagion and income heterogeneity in new product diffusion: a meta-analytic test". *Marketing Science*, 23(4), Fall 2004, 530-544.

Van den Bulte, C. 2000. "New product diffusion acceleration: measurement and analysis". *Marketing Science*, 19, 366-380.

All web links were live as of 23 September 2008.

Endnotes:

[i] The Assessment and Qualifications Alliance (AQA), Oxford Cambridge and RSA (OCR) Examinations, Edexcel, WJEC (formerly Welsh Joint Education Committee) and Council for the Curriculum, Examinations and Assessment (CCEA), which is based in Northern Ireland.

[ii] See: Mahajan, Muller & Bass (1990); Sultan, Farley & Lehman (1990); Parker (1994); Mahajan, Muller & Wind (2000); Bass (2004); Hauser, Tellis & Griffin (2006); Meade & Islam (2006); Muller, Peres & Mahajan (2007); Chandrasekaran & Tellis (2008) for reviews of this literature strand.

[iii] Bass model applications have now been used for many products over and beyond consumer durables – see below.

[iv] The formula is operationalised on in a spreadsheet available on the 'Bass Basement' website (http://www.frankmbass.org/).

[v] For further description of the types of social contagion in Van den Bulte & Stremersch's analysis, see below.

[vi] 'the extent to which the less powerful members of [a culture] expect and accept that culture is distributed unequally' (Van den Bulte & Stremersch, 2004, p.533)

[vii] This term can be taken to be synonymous with the term ‘network externalities’, as described above.

[viii] Shadow diffusion has also been referred to as the 'sleeper' or 'cult' effect – for example when a movie, book, piece of music, etc. becomes popular by informal, 'invisible' channels before formal, visible diffusion is measured.

[ix] These data were kindly provided by Peter Hermans of the Dutch national institute for educational measurement, CITO.