

Improving the Quality of G&T Nominations: Tackling Creativity

Dr Tom Balchin

Brunel University, UK.

Abstract

This paper considers some of the problems that occur when school teachers wish to convey the standard of individual students' creative outcomes to G&T co-ordinators, but lack specific assessment tools to record and transmit evidence. This demand was highlighted by a survey of the views of 800 G&T co-ordinators in 2005. The development of a new tool is disclosed, with which subject teachers can consensually assess the creativity of products with a high degree of reliability. Some formative results from its trialling revealed significantly high correlations of agreement (ranging from $r=.89$ to $r=.95$) over seven essential criteria, even when the different processes, climates and students themselves were unknown to the judging panels. The tool's effectiveness as part of a Creative Feedback Package (CFP) for teachers, tasked with nominating the gifted, is evaluated.

The demand to place value upon creative effort

Academics and practitioners who seek to improve the way education for the gifted is managed in Britain, need without doubt to concentrate on producing ways for teachers to understand how to accurately attach the label of 'gifted' to the various students under their temporary charge. The paper on pages ? to ? of this Journal explains how G&T co-ordinators who are tasked to nominate gifted students understand creativity to be a psychological construct inside what they think of as giftedness. When asked what signs they typically looked for in students to identify giftedness, a majority of the sample believed that indications of creative achievement were the one of the best way to identify giftedness. 'Creativity' was self-identified, not actively sought for. Descriptors of creativity returned were, or were variations of, the following: risk-taking, unusual approaches to problem solving, lateral thinking, the application of ideas or concepts to unfamiliar contexts, thinking outside the box or originality.

Co-ordinators felt that because creativity is such a subjectively-used term, they were unable to justify their opinions concerning individuals' creativity, or give credit for failure which results whilst attempting to bring original ideas into fruition without using some kind of appropriately validated test or tool. One in 8 of each school sampled actually used indications of creative thinking/effort as a guide to identifying giftedness... but did not possess any kinds of instruments or tools to evaluate it. A majority of the primary school teachers recorded that they were waiting for appropriate creativity tests to become available for their schools and did not feel able to adapt any other kinds of creativity assessments for outcomes to their own purposes with any kind of authority. It is clear that there is a strong need to produce tools to help co-ordinators justify their decisions concerning what is creative, and what is not.

There exists a powerful force in teachers' favour which can be capitalised upon here. The researcher has found persuasive evidence in over 40 classroom evaluations of the particular tool described in this study, that teachers can usually spot creativity fairly accurately when they see it. They do this by comparing outcomes with their own frames of reference to what they already know to be creative within their particular fields. This goes part way to explaining the high incidence of G&T co-ordinators sampled who specified that they used creativity as an indicator to giftedness.

The same teachers would then admit that they were not good at explaining their decisions in this regard, and relied on simply declaring that their professional judgement and intuition is valid. This is currently quite acceptable in British schools, as judgements of creativity are seen as highly subjective anyway. This factor alone is responsible for the failure of many well intentioned attempts to create appropriate holistic assessment systems over the years. However, it will be seen in the research described in this paper that this ability for teachers (as experts in their particular fields) to recognise creative effort, can be harnessed in order to carry out successful consensual assessments of creativity.

Some may argue that creativity is the last thing that should be assessed: as by its nature it would suffer under assessment. It involves taking risks – things that students tend not to do under rigid assessment parameters – which certainly exist in the kind of 'control culture' that it must be acknowledged exists currently in even the best of state schools.

Furthermore, teachers under pressure to meet targets can often consciously or unconsciously discourage creativity. Many G&T co-ordinators in the sample reported that their students tend to stick to what they know will definitely get them good marks, and reject, understandably, the possibility of failure by 'going out on a limb.' Risk-taking can be attempted, but if work simply has not been finished in time, or is badly done, it can simply fail assessment boundaries.

Creative, and possibly therefore gifted, children can go 'under the radar' and not be nominated by teachers, especially in primary schools. Their creative potential will not be picked up co-ordinators who are looking at portfolios or past work to look for signs of giftedness. It is not an easy task, especially when newly released statistics (Mansell 2007), show how non-tested subjects (typically the very subjects which allow the most scope for creative experimentation and expression) have been cut back drastically over the last 5 years due to the pressures of high-stakes testing in English, maths and science. PE, art, design and technology and music are typically affected, and this cannot help teachers tasked with nominating gifted students.

These important figures, from a 2003 monitoring survey carried out by the Qualifications and Curriculum Authority (QCA) are the most definitive ever provided on the extent of test coaching for Sats, but have only now been released under the Freedom of Information Act. They show that during January and May 2003, 44% of all lessons during the final year of primary school were devoted to test preparation and coaching; almost as much effort as preparation for GSCE's.

This does not help G&T co-ordinators at all, rendering accurate identification, especially of creativity gifted children, very tricky. Indeed, the 2005 survey of the co-ordinators revealed that they had to knowingly rely on the guesses of colleagues tasked to find certain percentages of students to be called labelled as gifted. Mansell (2007), states that teachers are taking short-cuts to improve the statistics, at the expense of pupil's true understanding and learning. It can be seen that this statement is reflected by the findings from the 2005 survey.

As children enter secondary schools, there is a natural 'discontinuity in progression' (Stables 1994) which describes the lowering of overall creative production due to having to learn specific skills and techniques. It can be seen that creative work is restricted in favour of learning knowledge (to pass Sats tests) now at the end of primary school as well as the first year of secondary school. The 2005 survey of British G&T co-ordinators highlighted how many gifted children in secondary schools tend to not to care about the final productions of their creative ideas, perhaps a reflection of the above. This is not helpful for the co-ordinators who reported that they are waiting for a way to judge creativity in unfinished products, otherwise they felt they could not justify their nominations for school G&T registers.

Poor work, or unfinished work, often serves to act as a misguided buffer against negative peer group pressures. The sample of G&T co-ordinators indicated that this factor is one of the biggest influences on the misidentification of gifted British children. However, it is speculated that once gifted students know their creative efforts are being supported and actively taken notice of as part of marking requirements, they may start to feel it is worth taking care to produce work that is creative, *and* well presented.

Perhaps a good start is to regard the information produced from assessments (particularly in the arts subjects) to be part of the solution to facilitate creative cultures, as well as being an information source for G&T co-ordinator for the identification of creatively gifted children. Further to this argument, if it can be made clear to curriculum planners that creativity *can* accurately be judged (even if it is not directly assessable and comparable to the results of other schools), then arts subjects (wherein many gifted children shine) can retain their importance, especially at primary school level.

As a preface to the research study, neurological data is discussed which seems to confirm the researcher's experience; that creativity actually demand a higher than average level of intelligence, and that highly creative children are likely to be gifted.

The creative brain

The process of man's creation, according to Crutchfield (1961), is not something inherently mysterious or unanalysable; but rather it is like any psychological process, subject to scientific research and analysis, to experimental manipulation and control. 'Creativity', a convenient summary label for a complex set of cognitive and motivational processes in an individual's brain, means that the main relevant research fields understand different definitions. For instance, a neuroscientist regards the creative process as the basic, still unknown set of internal neuro-physiological and neuro-chemical events which determine creative activity, and go along with it (Heilman, Nadeau, and Beversdorf, 2003). A psychologist on the other hand regards the process as a series of mental stages, steps, actions and behaviour, (for which there are a number of known strategies and quantities) which culminate in a novel idea or creative product.

A suitable definition of creativity, as a start-point, may be understood as the ability to understand and express novel orderly relationships. (Boden, 1990). There are too many gaps in our current knowledge of the brain in order to comment accurately on the physiological causes of the series of mental events that lead to creative outcomes, but the majority of research in this area indicates that it must emerge out of some sort of fusion of percepts, images, motor adjustments and concepts, which are in a constant flow of electro-chemical interaction and motor imagery.

Early researcher Rugg (1963) proposed that some autonomous forming process sweeps like a magnet across the chaotic elements of mind, picks up the significant segments and in a welding flash, precipitates the creative response. Neurologists Heilman, Nadeau, and Beversdorf (2003) now confirm that finding this thread might require the binding of different forms of knowledge, stored in separate cortical modules that have not been previously associated. Such forms of knowledge which interact with the stimuli might be divergent (and convergent) thinking, high level of general intelligence, domain-specific knowledge, and other special motor skills. The term 'multi-contextual thought' may more accurately express the idea that more than two planes of thought are involved in this type of thinking and expression.

Most hemisphericity studies (like those reported by Davis and Al-Sabaty in 1989), promote creativity as predominantly a 'right-brain' activity. Investigators Torrance and Taggart proposed back in 1984 that creative thought requires 'left-brain', critical, evaluatory types of thinking to continue ideation towards creative solutions. The first

neuro-imaging investigations of human responses to fluid analogies (Geake and Hansen, 2005) have produced evidence that creative innovation requires the immediate co-activation and communication between these regions of the brain that ordinarily are not strongly connected.

Furthermore, these imaging studies seem to promote the necessity of a strong dynamic between the individual's existing capability for cognitive representations of the problem (which affords mental access to the implicit relationships involved) and the individual's intelligence... because the more mental relationships that are evoked by an element, the more likely it is that another element will be combined with it in a manageable form. Associative hierarchies resemble generalisation gradients of different degrees of steepness, with associations to words or problems ranging from common to unique. Individuals with relatively steep gradients tend to give common associations at high strength, but few or no uncommon associations. In extreme, their responses may be rigid and stereotyped.

The basis of the Torrance tests is to elicit difficulties in order to prompt creative response. For example, an individual with a steeper gradient may respond with the more obvious: 'to hold pieces of paper together', rather than 'to pick a lock' or 'open a stuck CD-drive on a laptop' when asked to detail possible uses for a paperclip. To activate this type of relational processing, Gabora (2000) and Csikszentmihayli (1996) both recommend trying to use unusual connections and links in order to think around the nature of a problem.

There are two main (but conflicting) views of how the brain is creatively intelligent, both utilising associations of diverse elements. In the first, a precondition for creative performance is a clearing of conventional mind-sets. Urban (2003) reviews suggestions that this can occur under conditions of low cortical arousal, where defocused attention and flat associational hierarchies might maximise inter-modular activations. Since inhibition or suppression (by anxiety or other more powerful competitive stimuli) would limit awareness and open-ness to both internal and external stimuli, their loosening would favour associative thinking, and so creativity.

Latent inhibition, relatively depressed in creatively intelligent individuals (Carson, Peterson & Higgins, 2003) may enable the holding available, rather than rejecting *a priori*, of a greater number of solution trajectories to a problem. Andreasen (2005) states that her neuro-scientific research has produced evidence indicating that creative people may slip into a 'zone' in which ideas and thoughts come up freely in a disorganized way. She has found that during this state, a part of the brain known as the association cortex becomes very active. That brain region is known to be able to link up ideas or thoughts in potentially novel ways. It is capable of effectively processing increased inputs without the risk of cognitive overload.

In the second (more recently developed) view, researchers such as Geake (2004) and Hofstadter (2001) challenge proposals of low-cortical arousal states for maximal creativity, and propose that creativity actually requires an explicit manipulation of higher-order mind-sets. Geake (2004) believes that the cognitive basis of intelligence itself is the ability to make fluid or creative analogical associations between distantly related concepts or pieces of information.

Geake & Hansen (2005) have carried out the first (worldwide) testing, using fMRI scanning at Oxford University, UK., of subjects being creative. Their studies have produced evidence to suggest that it is less about dissociate states, but the possession of larger than normal working memory capacities that are responsible for enabling them to create more cognitive variance. This allows for possible outcomes to be held on-line for longer in order to weight their potential creativity (Geake & Hansen, 2005). This kind of ideational fluency and a preference for complex and asymmetrical designs (two of the main factors contributing to creativity) could, in fact (according to Geake and Hansen, 2005) derive from higher levels of arousal and from faster stimulation of discrete cerebral areas. Put simply, they have found evidence to indicate that more intelligent (and prepared) people are also the most creative.

Although in UK schools at the moment, at least, it seems *de rigueur* for the school prospectus or website to pronounce that each child is equally capable of being as creative as another at any one time, it seems that the most creative individuals are capable of a widening of selective attention, which renders them more receptive to learning experiences. Furthermore, this probably comes with a more intensive sampling of environmental stimuli.

The researcher can tentatively add some conclusions from minute by minute observations of 12 classes of students in schools around the UK (Balchin, 2005a) carrying out extended (two-day) design and technology tasks. Time and time again, it was recorded that the 'drivers' of the small groups working on design and make tasks, were the students with the highest grades in other subjects and the highest degree of motivation relative to their peers (according to their teachers).

By giving them 'creative moments sheets' to complete, a strong link was found between the students' successes with conquering design and make problems in creative ways, and their enthusiasm for the recall of related learning experiences, which they could link to the problem in order to propose a range of solutions (and then to zero in on the most appropriate). The most successful were able to draw on their *understanding* of those experiences and relate it in some relevant way to the task confronting them.

The model below brings a neurological understanding to the most recently proposed (business-driven) cyclical-type models of creative ideation, which are usually characterised by dynamic and chaotic iterations between 'mind and hand' (eg. Plsek, 1997). Motivation can be seen to play a crucial part, and evaluation is continual. It seems to indicate that the memory functions as the neural 'starting line' from which creative outcomes emerge, depending on the countless ways that the 4 P's are aligned

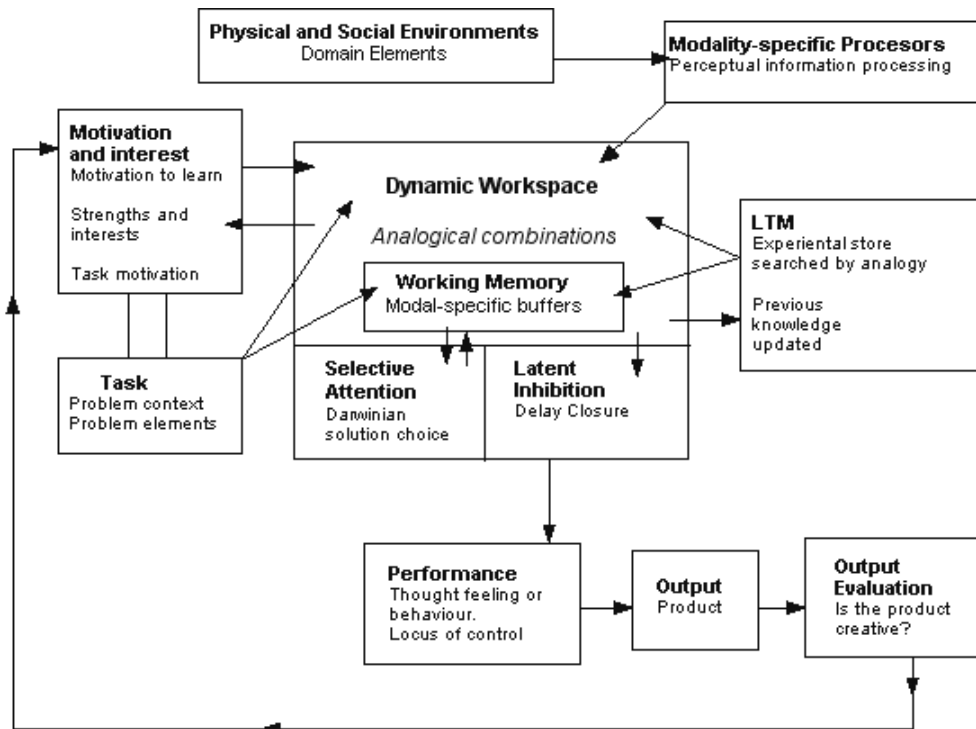


Figure 1. Geake and Dodson's 2005 neuro-psychological model of the creative intelligence of gifted children.

The ability to associate elements is argued by the model above to be a dimension of intelligence. Teaching to maximise a student's intellectual abilities, then, should include teaching to maximise a student's creative intelligence through the challenge of high-quality creative output. It is now possible, seemingly, via neuro-imaging, to view the creative process... but Cartier (2000) and Reis and Renzulli (1991) suggest that it is logical to presume that without a creative product, we cannot conclude that a person is creative. Cropley (2006) explains that the kind of free production of variability, through unfettered divergent thinking, holds out the seductive promise of effortless creativity, but it actually runs the risk of generating only quasi-creativity or pseudo-creativity if it is not adapted to reality.

The Creativity Feedback Package

Creativity is considered to be a third-part of giftedness (Renzulli 1986), but a big question therefore for teachers is attaching value or worth to creative thinking which any sort of creative result or production. Given the strong bias towards the necessity for production from creative activity it is little wonder that much effort has been devoted over the last fifty years, mainly by US researchers, to the identification of the special characteristics that lead to creativity, to the ways that creative people solve problems, and to producing criteria to conclude the existence of a truly creative outcome.

These efforts have led to the development of numerous instruments that are used in the identification of creative talent. Kaltsounis and Honeywell (1980), for example, have published an exhaustive list of 'creativity tests' and Torrance and Goff (1990) have identified 255 such instruments. It must be emphasised that different researchers

have studied different aspects of creativity and that often these cross-purposes have been reflected in the measuring processes. It seems that by using distinct techniques (tests of cognition, attitudes, interests, personality, biography, etc) to assess creativity, researchers are actually studying different phenomena. Conclusions with a fair degree of validity, in consequence, are limited to studies incorporating the same method. Theory integration difficulties mean that currently, there does not exist one accepted method for the measurement of creativity in individuals or groups working collectively.

The Creative Feedback Package (CFP) is a new technique which has been developed by the researcher as part of an effort to help with the creativity area of G&T identification. The package is being developed now in conjunction with other tools for the British education market and recommended for use by nominating teachers who wish to provide evidence to co-ordinators of creative activity. The Package holds a tool for the creative process, called the 'Creative Moments sheet' and a tool for judging the effectiveness of the environment for creativity, called the 'Creative Climate sheet'.

The Creative Product feedback sheet, discussed below, has just seven criteria of creativity for products to be scored against. These were generated in 2006 by constructing lists of descriptors of creative products from the extensive literature (mainly US produced) then repeatedly subsuming the available criteria until only the very essentials emerged. It can be seen that the form derives its base format in particular from the Creative Product Analysis Matrix (CPAM) theory produced by Besemer and Treffinger in 1981. The CPAM is a three-dimensional model of creativity in products, which hypothesizes novelty, resolution, and elaboration/synthesis as three factors which most help to judges to focus their attention on relevant attributes of products (Besemer 1998). The form was then trialled in schools around Britain for nine months in order to see if could retain fully the complexity of the phenomenon.

Four criteria describe the creative concept, or ideas, and three criteria describe the quality of outcome; which evaluates how well the creative thoughts have been shown in the product. It is proposed that creativity is seen in both the concept and the standard of production that the result showed. But it is the creative concept stage where the unique ideas can be seen to be brought forth, and the quality of outcome stage is the manifestation of these. The latter cannot occur without the former; it is proposed that the quality of outcome is a 'vehicle' for the creative thought. The particular format for assessing products made in design and technology classes is shown below:

Uniqueness - what is the degree of deviance from the normal incidence?

Association of ideas - is there a link or attempt at synthesis of two or more broad ideas?

Risk-taking - is it a bold attempt? How ambitious it? Is there a clear challenge taken on?

Potential - does it have the capacity to succeed/solve the problem/what is its likely promise?

Operability - is it effective? Does it really work?

Well-craftedness - is it carefully/neatly/robustly made?

Attractiveness - is it enjoyable/surprising/pleasing to look at?

The 'association of ideas' in particular, as seen on Geake and Dodson's 2005 model, is a most important criterion at the heart of many of the accepted overall definitions of

creativity, and it was important to find out whether simply asking judging panels whether they could tell such information from a product alone, would work. During trials, clear evidence emerged to show that it did work. Checking with information on processes from the students involved showed that the products nearly always revealed strong clues to the inspiring factors involved - enough for the judges to feel confident about reaching conclusions. Clearly, knowledge of the task, processes and skills levels of the creator assists this confidence.

In general, if the condition of the item is poor, creativity is harder to see and score. This is why operability, well-craftedness and attractiveness are important markers in creativity judgements, because they relate to the feeling of rightness of a product and the sense that ideas have been pulled together into a coherent whole.

The sheet incorporates an adapted Likert scale (with 12 points) to use for each criterion; intended to mirror the A, B, C and D of traditional marking schemes, with the corresponding pluses and minuses. This scale has been shown to have the same effect of forcing opinions as a 4 point scale, as the marker cannot 'sit in the middle'. Importantly, measures taken from the scale can either be looked upon as having value only in the way they help to force judgements by getting the scorers to focus their attention on the criteria, or to produce concrete measures. The version being developed for 2008 focuses on the production of a single score by a small panel of teachers, who can total-up the scores from the scale and divide by the number of criteria to produce an overall score for an individual's or group's creative achievement

Consensual assessment

This concerns the idea that a product or response is creative to the extent that appropriate observers consensually agree it is creative. Appropriate observers are those familiar with the domain in which the product was created or the response articulated. This is not a particularly profound proposition... but the technique has been shown to be more reliable than most other tests, probably because it involves the views of experts from the domain the creative work under examination resides in. Teachers, who can be considered as experts in a subject domain, can recognise how a contribution from a student may vary from established works, constructs or the thinking or performances of other students.

The most important criterion for creativity assessment procedures is that any ratings produced should be found to be as reliable as possible (Cropley 2000). By definition, inter-judge reliability in this method is equivalent to construct validity; if appropriate judges independently agree that a given product is highly creative, then it can and must be accepted as such. It should also be possible to separate judgments of creative conceptualisations from judgments of technical goodness and aesthetic appeal. The reliability of this technique in relation to others, is consistently high. For instance, Amabile's formative studies in 1982 show high levels of agreement (correlated between .73 and .93 agreement) between raters judging the creativity of collages. Her technique requires expert judges to rate the creativity of a product on a scale of 1 to 5 using a subjective definition instead of a given criterion or checklist.

Baer, Kaufmann and Gentile (2004) tell us that its validation, however, has been limited to assessing the creativity of artefacts produced under tightly constrained

experimental conditions. Typically, only those produced in response to very similar instructions have been compared. This has allowed researchers to compare such things as the effects of different motivational conditions on creative performance, but has not allowed many other kinds of comparisons. It has also limited the use of the technique to artefacts gathered for specific experimental purposes, as opposed to already-existing artefacts produced under less controlled conditions.

A 'health warning' exists (suggested by Hocevar and Bachelor, 1989) for groups who gather to perform consensual assessments of creativity. A dangerous feature of small groups is that it is possible that they can establish some particular opinion-sets from (or before) the outset, which polarise their judgments. Beckwith and Lehmann (1975) suggested that this may be termed the 'halo effect'. Surowiecki (2005) found that groups (supplied with knowledge about a situation or product with a similar few pieces of knowledge in common) tend to discuss only the information in common that they view as most important; which may impact negatively upon consensual assessments.

This body of research and opinion seems to argue the case for the relative ineffectiveness of small introverted groups making decisions. But, as solid as the evidence for such polarisation effects are, so too is the evidence demonstrating that that small groups actually make better decisions and come up with better answers than most of their members. The group decision outcomes may outperform even its best member (Sunstein, 2003).

It follows that there are some factors that must be considered when performing consensual assessments. The first problem is to decide who the judges should be. In school-level education it is the teachers, perhaps with the help of field-based professionals. It may also be useful for students to be allowed access to this group, in order for them to observe and gain understanding of the ways in which products are being evaluated for creativity. A result may be the initiation of dialogue amongst pupils about the way they can actively seek to make their ideas more creative, rather than just settling on the first solution that seems to solve the solution posed by a question or task brief.

The second question is to decide what the judges should be looking for. Not surprisingly, tutors may be defensive of their own pedagogies, and different researchers have asked their judges to consider different criteria for creativity. The list of criteria is virtually unending, including fluency, originality and other cognitive characteristics, personality traits, self-expression, enthusiasm, productivity, and expertise. Some investigators have used elaborate definitions of creativity; others use no definition at all. Furthermore, there is no guarantee that judges will understand and be guided by the sometimes complex definitions or instructions.

Method

Participants

A research study to evaluate a consensual assessment tool for products was performed in 2006 at Brunel University, London, UK, with 14 design and technology (hereafter d&t) teachers from schools around Britain (2 panels of 5, and 1 panels of 4) as 'judges'. Six teachers were women, 8 were men, mean age = 37 years. Volunteers were not paid for their participation; they were drawn from schools with strong links

to the University. The aim was to gather experts in a very practical educational field to act as judges of three dimensional products made by their students. The researcher held information concerning whether the creator had been identified as gifted.

Stimuli

The creative product feedback sheet was designed as a score card to guide the teachers in judging and marking a product for creativity. It was hoped that hitherto strangers could use the consensual assessment technique in an independent, and then discursive, manner. The discursive process used is not a weakness in the approach; inter-judge reliabilities reported in this area of research have been found to be high – in excess of .90 (Pearson product moment co-efficient), but does rely on a small group of teachers having the interest and time to do it. Numbers were not critical to this particular exercise, because it was decided that the judge will inevitably have his/her unique scale of judgement and frames of reference to the product. This study would test the product sheet and the technique, and not focus on how scores could be produced in order to aid identifications of the gifted. The scoring of each criteria therefore was more about forcing debate and consensual agreement.

Procedure

14 products were brought along, ranging from a full-size adjustable garden chair to a necklace made from a variety of metals. The judges were all given the individual product assessment sheet, asked to choose any 2 of the 14 products they liked, and given 30 minutes to assess them according to the criteria above. The one rule given was that if more than one were working on the same product, no collaboration would be allowed. It is important to bear in mind that no accompanying portfolio/coursework was present, nor did the judges have any information about the students to concerning whether they had been identified as gifted or not. Therefore, the judgements were made from the products alone, to see if agreement over the criteria could occur even when teachers did not know anything about the task or making process.

This had the effect of getting the raters used to the questions inherent in the criteria, as well as providing comparative data. After this individual scoring time, the 14 judges were asked to split into 3 groups and given 40 minutes to 'rotate around and assess' 3 out of the 14 products. The products chosen for investigation were a garden chair, an aluminium lamp and a hat. They were tasked to debate and reach a consensus before scoring each criterion on the 1-12 scale. Serious debate was seen to take place inter-group. However it was observed that the judges were quite able to agree about criteria, without having to compromise their initial thoughts too much. In this way, individual scores were produced for 14 products and consensual scores from 3 groups were produced for 3 products.

Results

The product chosen to represent the research here had been made by a Year 9 girl in a Durham school who had been identified as gifted. See this Journal, 9(1), 20-26 for an account of how the task was set-up to highlight indications of giftedness, and the comparing of her activities to the DfEE-produced indicators of giftedness in design and technology. We were told by the school that this girl was worth observing in a group situation to see how she would respond to the 2-day intensive design and make task. The result, a hat, is chosen here, because the full process of design and make had

been observed and recorded (but was not made available to judging panels). However, it was important that we had this information; the school viewed her as gifted, and the process of design and make that we observed convinced us that she was very creative. We speculated on whether creative students could be pronounced as gifted by default

The hat had also been scored by 3 teachers during the individual assessments and was the most complicated of the 3 products scored consensually, due to the girl trying to incorporate a mock ‘cooling element’ into the rim, and its partially unfinished nature. It represented the product which attained the least agreement between raters over the 7 criteria; in order to try and give a fair representation of the efficacy of the technique allied to the CFP’s product sheet.

The results are reproduced below (using those produced by evaluating one of the products as an example) for the seven criteria used, with the degree of spread between the scores given in each criteria marked with circles. The horizontal line that many are on indicates the lateral degree of spread of each of the scores between 1 (furthest left) and 12 (furthest right). The spread of scores are indicated in the column of numbers to the right.

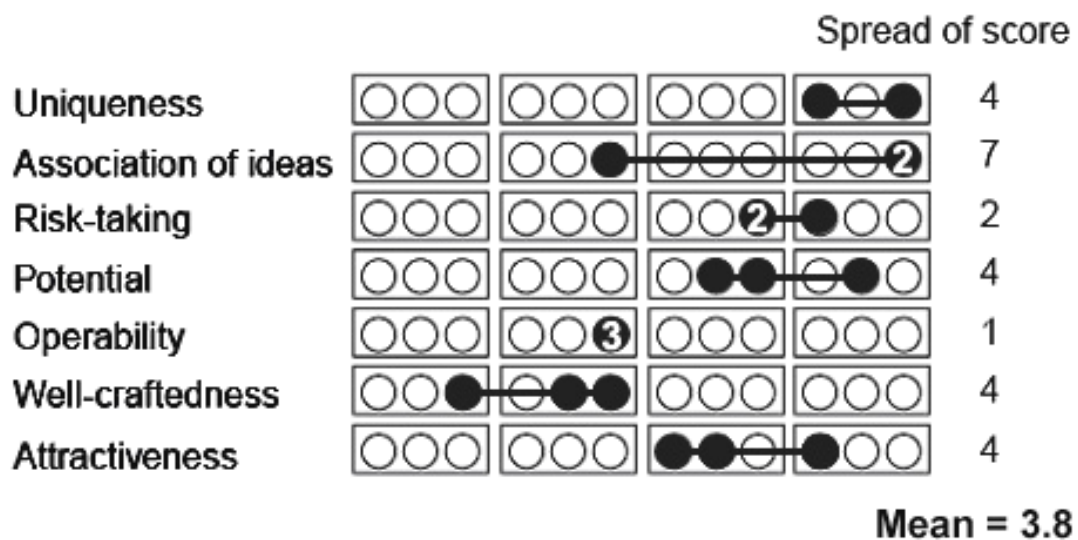


Figure 2. Spread of scores from three individual scorers of one product.

The table above shows the result (spread of scores) from 3 individual assessments of the hat. Only one of the seven criteria, the association of ideas, fell outside a spread of four marks. This was a difficult definition to present to teachers, but still, two out of three scored it in exactly the same way. The other was working from her particular frame of reference, which is a clue to the value of consensual assessing, which seems to have the power to take most prior frames of reference into account.

Taken as a whole, the judges (scoring individually) showed a fairly high degree of consensus over each score. When analysing the data and comparing rankings from the groups of teachers, correlations were used to compare them. This is expressed as the Pearson product moment correlations coefficient, which is a dimensionless index that ranges from -1.0 to 1.0 inclusive, and reflects the extent of a linear relationship

between the two data-sets. The overall correlation of agreement concerning the ‘creativity’ of the hat was revealed to be $r=.69$. 3 groups of 5 or 4 judges then looked at the same product and scored it consensually. Below is a table showing the results of their scoring.

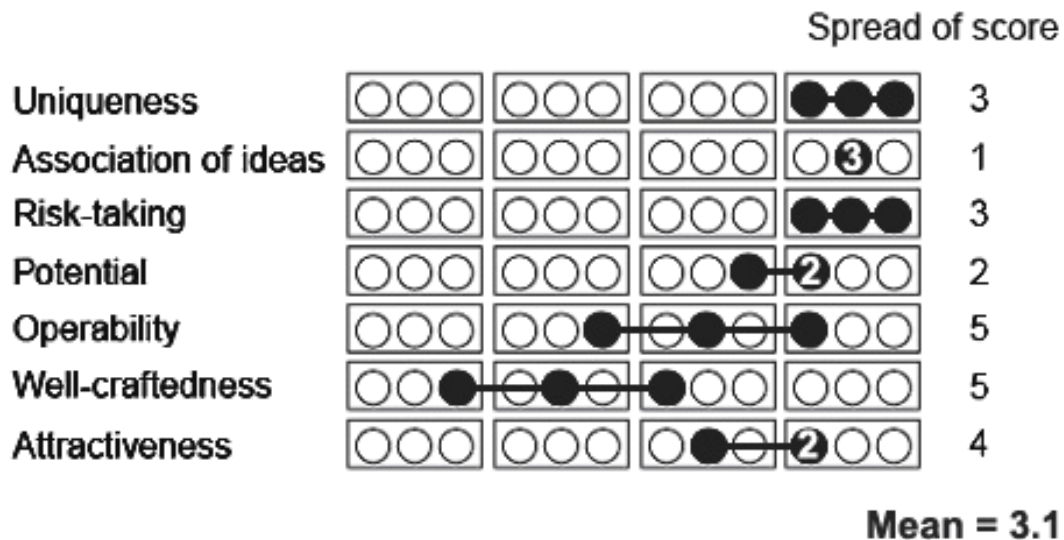


Figure 3. Spread of scores from three groups consensually scoring the product.

The product moment correlation was revealed to be $r=.90$, a significantly higher result from the products scored individually. An average of all scores for both assessments by 3 individuals and 3 individual groups was taken. The result: mean=3.8 (for the 3 individuals scoring the product) and mean=3.1 (for the 3 individual groups scoring the product)

Discussion

This result from the individual scoring seemed to indicate that (most of) the criteria and definitions worked, and that there was enough ‘clear space’ between each definition for each to make sense. It also showed that the judges were pretty good at recognising the inherent creativity of a product without any discussion, simply by reference to other similar products that they are aware of. Moreover, individuals could arrive at similar conclusions about the products without viewing any of the designing and making of the product.

As the products were divorced from the makers, did consensual scoring tend to diminish personal bias or dislike, giving a fairer impression of creativity inherent in the product? And crucially, could the creativity visible in the product imply a level of giftedness? It can be seen that the technique itself seems to work: when scoring consensually, it can be seen that the mean spread of scores per criteria diminished when the judges were allowed to sharpen and hone their professional judgements about the product through debate aimed at reaching agreement. An average of all scores for both assessments by 3 individuals and 3 individual groups was taken. The result confirms the background research, which indicates that assessing in small groups of judges (each experts from the same domain) serves to increase the reliability of the assessment in question.

The question of the tool's efficacy for the identification of giftedness is now discussed. The first four criteria concern the product's concept. All four were scored very close to each other. For example, when it came to scoring for the association of ideas, each consensual group of 5 or 4 judges scored it exactly the same (without conferring between groups, as per instructions for the whole exercise). This was a significant and positive finding, because the hat was unfinished. Nevertheless, the judges were able to see through the unfinished state and glimpse the creative thoughts that caused the construction to be the way it was.

The unanimous agreement about this particular criterion (produced by 3 groups of judges, each scoring consensually) held fast with the other 2 products under group assessment (the chair and the lamp). The association of ideas is a most important criteria at the heart of creativity, and it was important that each group of consensual assessors were able (with a one-line definition) to understand the notion well.

The last 3 criteria concerned the product's quality of make. In general, if the condition of the item is poor, creativity is harder to see and score. This is why operability, well-craftedness and attractiveness are important markers in creativity judgements, because they relate to the feeling of rightness of a product and the sense that ideas have been pulled together into a coherent whole. It can be seen that the three individual judges, and the three consensual judging groups especially, dispersed their scores for the product's quality of make around the scoring box, from 3-10. This was in contrast to the dispersal of scores for the product's concept; where the scores ranged from 8-12 (except for one score given to association of ideas by one group)

Upon asking the judges about this disparity, the answer came back was that the hat was difficult to judge in terms of operability because there were different opinions over the intentions of the students who made it. Some of the judges thought that the student who made it had hoped the blue liquid encased in acetate in the top of the hat would actually 'work' to cool down the head. This was purely concept modelling, but they had taken this into account whilst group scoring. This was an un-foreseen element but it did highlight some of the problems of groups of judges assessing without knowing what the students' tasks were all about. This study was designed partly to bring such elements to the surface, and to see how far groups can agree upon marks to be awarded for creative giftedness without knowing fully the accompanying context or process that an artefact has gone through during its building.

The judges seemed to be able to converge on scores to give to the concepts even though their ideas about the standard of build that should be achieved for the creativity adequately to be shown, differed. This was an indicator that creativity can be reliably recognised in a product, even if it was badly made and did not 'work'. This offers hope for a potential solution for teachers who feel they cannot recommend students to G&T co-ordinators for extra provision if they present badly made work. It may be remembered that one co-ordinator from the sample of 800 asked; 'how can I give them extra work, if they don't do the work I set them to do now?'

The judges therefore agreed that the creativity resides in the concepts or idea. They could see that a product which may score badly in terms of well-craftedness, operability and attractiveness, but still score very highly in terms of uniqueness, association of ideas, risk-taking and especially, potential. They believed that it could

encourage more creative thought during designing stages, and begin to give some sort of feedback/recognition to creative 'failure'. They particularly viewed these assessments as a good way of showing students that there are criteria inside creativity (mind or hand) that their products can pass or fail on – so that the assessments can be shown to them with advice to concentrate more on creative thinking, or presentation of work.

By analysing the remaining 2 creative product sheets, which recorded the consensual assessments of a table-lamp and a chair, it became clear that the same phenomenon of high correlations between raters had occurred. The result from the lamp (produced from 3 groups who did not discuss results between them) is an overall correlation of agreement between criteria ($r=.93$). The judging groups converged nearly exactly on each of the 4 aspects of the creative concept, showing that teachers can identify creativity. The evaluation of the chair showed extremely high correlates of agreement (again from 3 groups who did not discuss results between them). The raters did not judge the chair as particularly unique, but its potential was high. A very high overall correlation of agreement, $r=.95$ was obtained. It was interesting to note how as the products became less 'unique', the judging groups abilities to agree on product creativeness ratcheted up. This is due again to our frames of reference: we see chairs everyday, but we hardly ever see hats that aim to cool you down.

This was a relatively small sample of judges and products in comparison to the major studies using this technique (eg. Amabile, 1982), but it illustrates a process for developing and refining judgements about creativity that may add usefully to the growing body of research on consensual assessments of young people's ability. This study also indicates that a fewer number of indicators or criteria may strictly be necessary, accurately to judge facets of creativity within a product, than are currently present within Amabile's system of 20 product creativity ratings.

During this study, gathering scores was more for the purpose of feedback for the teacher, to gather information to filter down to G&T co-ordinators, rather than the scores themselves becoming the end result (Balchin, 2005b). Nevertheless, here is a technique which, once adapted to create simple marks for creativity in the manner suggested in the background research, can help teachers tasked with identifying gifted children.

Whilst an enthusiastic approach towards helping G&T co-ordinators to ramp-up the efficiency of identifications should be encouraged, any form of extra assessments inevitably create many negative questions, especially in the light of Mansell's recent findings about the intense pressure in schools caused by national testing (Mansell, 2007). The creative product sheet therefore needs to be looked at another way. By far the best feature of the tool was that it facilitated rich discussion about creativity and giftedness among the judges. This might then be the baseline knowledge for developing negotiations with students.

Shore, Rejskind and Kanevsky (2003) argue that there is convincing evidence that meta-cognitive training, such as getting students to verbalise their thinking, can be effective in enhancing children's perception of their own creative abilities. A way forward may be to provide a start-point for students by showing them the feedback from each criterion marked with these consensual assessments, in order to allow them

reflect on their own creative efforts, and perhaps construct new understandings of what their personal creativity means in the particular contexts in which it was utilised. The information can help not only the nominating teacher, be a useful feedback mechanism for the students.

This study found a high degree of agreement over the creativity of a product made by a girl who had been identified as gifted. Furthermore, a research study had been conducted which concluded that the product was very creative, even though it was unfinished. The judges were able to see this, knowing nothing about the product, and if we can accept that positive judgements of creative concept, whatever the presentation of outcome, is an indicator of giftedness, then the Creative Product sheet seems to work well. It seemed to show that there is no need for lengthy indicators of creativity in order to produce high correlations between small groups of judges who understand the contexts in which outcomes have been produced. Further trials have shown that even 2 teachers can produce the same high levels of agreement over the criteria.

It is hoped that this Creative Product sheet can, in the near future, be made part of a new downloadable package, designed to assist the production of more reliable nominations of students for provision programmes. It could also be adapted to situations where outcomes other than artefacts are sought. For instance, in the solution of problems in the performing arts, appropriate criteria may substitute for the 'quality of build' criteria. The prompters for discussion are easily adaptable to productions in other disciplinary fields.

References

- Amabile, T. M. (1982). Social psychology of creativity: a consensual assessment technique. *Journal of Personality and Social Psychology*, 43, 997-1013.
- Andreasen, N. (2005). *The creating brain: the neuroscience of genius*. Ezcurra, Mexico: Agrupacion Sierra Madre, 160-168.
- Baer, J., Kaufman, J. C., & Gentile, C. A. (2004). Extension of the consensual assessment technique to nonparallel creative products. *Creativity Research Journal*, 16(1), 113-117.
- Balchin, T. (2005a). A creativity feedback package for teachers and students of design and technology in the UK. *International Journal of Design and Technology Education*, 10(2), 31-43.
- Balchin, T. (2005b). Identifying gifted and talented students through creative behaviour. *The Journal of the National Association for Gifted Children*, 9(1), 20-26.
- Beckwith, N., & Lehmann, D. (1975). The importance of halo effects in multi-attribute attitude models. *Journal of Marketing Research*, 12(3), 265-275.
- Belcher, T. L., Rubovits, J. J., & DiMeo, P. A. (1981). Measurement of creativity: A factor analysis study. *Psychological Reports*, 48, 819-825.

Besemer, S. P. (1998). Creative Product Analysis Matrix: Testing the model structure and a comparison among products - Three novel chairs. *Creativity Research Journal* 11(4), 333-346.

Besemer, S. P. & Treffinger, D. H. (1981). Analysis of creative products: review and synthesis. *Journal of Creative Behavior* 15(3),158-178.

Boden, M.A. (1990). *The creative minds : myths and mechanisms*. London : Weidenfeld and Nicolson, 5-89.

Carson, S. H., Peterson, J. B. & Higgins, D. M. (2003). Decreased latent inhibition is associated with increased creative achievement in high-functioning individuals. *Journal of Personality and Social Psychology*, 85(3), 499-506.

Cartier, F. (2000). Review of 'Understanding those who create' *Mensa Journal*, (46)1, 79.

Cropley, A.J. (2000). Defining and measuring creativity: are creativity tests worth using? *Roeper Review*, 23(2), 72-77.

Cropley, A.J. (2006). In praise of convergent thinking. *Creativity Research Journal*, 18(3), 391-404.

Crutchfield, R.S. (1961). The creative process, in D.W. McKinnon et al. (Eds.), *Conference on the Creative Person* (pp. 58-93). Berkley, California: University of California Press.

Csikszentmihayli, M. (1996). *Creativity, flow and the psychology of discovery and innovation*. New York: Harper Perennial..

Davis, G. A., & Al-Sabaty, I. (1989). Relationship between creativity and right, left integrated thinking styles. *Creativity Research Journal*, 2(1), 111-117.

Gabora, L. (2000). The beer can theory of creativity in P. Bentley & D. Corne, (Eds.) *Creative Evolutionary Systems*. London: Morgan Kauffman.

Geake, J. G. (2004). How children's brains think: not left or right but both together. *Journal of Education 3-13*, 32(3), 65-72.

Geake, J. G. & Dodson, C. S. (2005). A neuro-psychological model of the creative intelligence of gifted children. *Gifted & Talented International*, 20(1), 4-16.

Geake, J. G., & Hansen, P. (2005). Neural correlates of intelligence as revealed by fMRI of fluid analogies, *NeuroImage*, 26(2), 555-564.

Halford, G. S. (1992). Analogical reasoning and conceptual complexity in cognitive development. *Human Development*, 35(4), 193-217.

Heilman, K. M., Nadeau, S. E., & Beversdorf, D. O. (2003). Creative innovation: possible brain mechanisms. *Neurocase*, 9, 369-379.

Hocevar, D., & Bachelor, P. (1989). A taxonomy and critique of measurements used in the study of creativity, in J.A. Glover, R.R. Ronning and C.R. Reynolds, (Eds.), *Handbook of creativity* (pp. 53-75). New York: Plenum Press.

Hofstadter, D. R. (1995). *Fluid concepts and creative analogies*. Basic Books, New York.

Holland, J. L. (1959). Some limitations of teacher ratings as predictors of creativity. *Journal of Educational Psychology*, 50, 219-223.

Isaksen, S. G. (1987) *Frontiers of Creativity Research: Beyond the Basics*. Buffalo, N.Y.: Bearly Limited, 341-379.

Kaltsounis, B., & Honeywell, L. (1980). Instruments useful in studying creative behaviour and creative talent. *Journal of Creative Behaviour*, 14, 56-67.

Mansell, W. (2007). *Education by Numbers: The Tyranny of Testing*. London: Politico's Publishing Ltd.

Murdock, M. C., & Puccio, G. J. (1993). A contextual organiser for conducting creativity research, in S.G. Isaksen, M.C. Murdock, R.L. Firestien, & D.J. Treffinger (Eds.), *Understanding and recognising creativity: the emergence of a discipline* (pp. 249-280). Norwood, N.J.: Ablex.

Plsek, P. E. (1997). *Creativity, innovations and quality*. Ontario: Irwin. 47-58.

Reis, S. M., & Renzulli, J. S. (1991). The assessment of creative products in programs for gifted and talented students. *Gifted Child Quarterly*, 35(3), 128-134.

Renzulli, J. (1986). The three-ring conception of giftedness: a developmental model for creative productivity, in Sternburg, R.J. and Davidson, J.E. (Eds.) *Conceptions of giftedness*. New York: Cambridge University Press.

Rhodes, M.(1961). An analysis of creativity. *Phi Delta Kappan*, 42, 305-310.

Shore, B. M., Rejskind, G., & Kanevsky, L. S. (2003). Cognitive research on giftedness: a window on creativity and teaching for creativity, in D. Ambrose, L. M. Cohen & A. J. Tannenbaum (Eds.) *Creative intelligence: toward theoretic integration* (pp. 181-210). Cresskill, N.J.: Hampton Press.

Stables, K. (1994). Discontinuity in transition: Pupils' experience of technology in year 6 and year 7. *International Journal of Technology and Design Education*, 5(2), 157-169.

Sunstein, C. (2003). *Why societies need dissent*. Cambridge: Harvard University Press, 32-70.

Surowiecki, J. (2005). *The wisdom of crowds: why the many are smarter than the few*. London: Abacus, 173-189.

Sternburg, R. J., & Lubart, T. I. (1991). *An investment theory of creativity and its development*. *Human Development*, 34, 1-31.

Torrance, E. P. & Goff, K. (1990). Fostering academic creativity in gifted students. Retrieved 18th December 2006, from Educational Resources Information Centre (ERIC) database, ED321489, ERIC Digest No. E484.

Taggart, W. & Torrance, E. P. (1984). *Human information processing survey, administrator's manual*. Bensenville, IL: Scholastic Testing Service.

Treffinger, D. J. (1987). Research on creativity assessment, in S.G. Isaksen (Ed.), *Frontiers of creativity research: beyond the basics*. Buffalo, N.Y.: Bearly Limited, 103-119.

Urban, K. (2003). Toward a componential model of creativity, in Ambrose, D., Cohen, L. M., & Tannenbaum, A.J. (Eds.) *Creative Intelligence: Toward Theoretic Integration*. Hampton Press Inc: Cresskill, N.J.