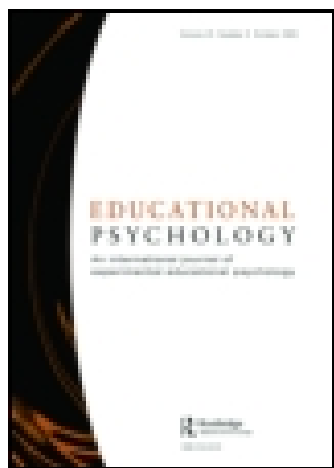




## Educational Psychology: An International Journal of Experimental Educational Psychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/cedp20>

### Delphi: a technique to harness expert opinion for critical decision-making tasks in education

Mark J. Clayton<sup>a</sup>

<sup>a</sup> Special Education Centre , Macquarie University , Sydney, Australia

Published online: 29 Sep 2006.

To cite this article: Mark J. Clayton (1997) Delphi: a technique to harness expert opinion for critical decision-making tasks in education, Educational Psychology: An International Journal of Experimental Educational Psychology, 17:4, 373-386, DOI: [10.1080/0144341970170401](https://doi.org/10.1080/0144341970170401)

To link to this article: <http://dx.doi.org/10.1080/0144341970170401>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms



## *Delphi: a technique to harness expert opinion for critical decision-making tasks in education*

**MARK J. CLAYTON**, *Special Education Centre, Macquarie University, Sydney, Australia*

**ABSTRACT** *Expert rather than general or informal opinion is often sought in the development of educational policy. Decisions to be made regarding best practice, the most effective way to deliver services, issues dealing with the professional development of teachers and the distribution of limited educational resources are examples which require critical thinking and reasoning. Regardless of the nature of the task, complex decision-making is rarely left to the remit of one person and there is usually an assumption made that 'two heads are better than one'. The organisational requirements of collecting, analysing, refining and validating critical information can be a long, arduous and often tedious process—a process which can often be overlooked, resulting in ill-defined, poorly conceived, biased and invalid determinations. The conventional Delphi procedure offers decision-makers a user-friendly, rigorous and systematic strategy in the collection and dissemination of critical information. This paper reviews the substantive literature relating to the Delphi procedure, provides a rationale for its use, describes the distinctive features, reviews key points of contention and provides an indication of both past and present uses.*

How are systemic educational objectives arrived at? Are the programmes for regular education different from special education? How should teachers be trained? Should all teachers-in-training be taught special education instructional skills? What is the best way to deliver professional development to teachers in rural areas? What are the teaching competencies required of today's teachers? Do teachers in rural areas require different skills from their colleagues teaching in metropolitan areas? Do teacher training programmes meet the immediate and/or future needs of their graduates' employers? These are questions of substance and are examples of the issues which policy-makers, politicians, bureaucrats and university academics are required to consider, debate and find solutions to. The process by which significant stakeholders follow in the pursuit of

answers to these questions is critical to the success, validity and credibility of the ideas and solutions generated.

The purpose of this paper is to inform researchers and others wishing to consider the expert judgements of others, of a systematic, rigorous and effective methodology designed to elicit potent and valid user-friendly answers to such questions. This paper will provide a review of the substantive literature relating to the Delphi procedure, discuss its distinctive features and provide an indication of both past and present uses in an attempt to whet the appetites of those embarking on critical decision-making exercises with particular reference to education.

For those decisions which can be made by flipping a coin, the decisions may not really matter; however, the effects of critical decisions may linger and when a mistake is made, the damage may be irreparable and extremely costly. This is why critical decisions, the kind involving personnel, programme improvement and management, and resource allocation, for example, require accurate information, careful consideration and involvement of more than a single decision-maker (Rasp, 1973).

Inevitably the person charged with the responsibility of securing greater involvement frequently employs either one or all of three techniques: single expert, several experts and round-table consensus. All three leave room for improvement: the single expert method because one person's judgement is simply not sufficient input; the several experts method because the individuals consulted have neither the opportunity to provide their most thoughtful input nor the benefit of hearing other responses that might encourage a refinement of their contributions; the round-table consensus because group decisions reflect the special characteristics of group dynamics and their potential distortions more often than objective truth.

In considering the latter technique, the round-table consensus, it is worth considering Forsyth (1990, p. 308) who maintains that "groups don't urge restraint; instead, they polarize opinions" when referring to the advisory group President Kennedy formed during the Bay of Pigs crisis. According to Forsyth, Kennedy probably assumed that a group "if faced with a choice between a risky alternative (such as invade Cuba) and a more moderate alternative (such as use diplomatic means to influence Cuba) would prefer the moderate route" (p. 308). Hindsight informs us however, that "groups' decisions actually tend to be more extreme than individuals' decisions".

This phenomenon dubbed 'risky-shift' was well researched during the decade from 1960 to 1970. Numerous studies revealed that risky-shift was a phenomenon of group discussion which seemed to intensify all sorts of attitudes, beliefs, values, judgements and perceptions. Sometimes referred to as 'choice shift' (see, e.g., Brandstatter *et al.*, 1982), risky-shift assumes that individual group members are persuaded to change their personal choices within a set of various response alternatives. The effect is shown when the group response is observed to be different (or risky) from the mean pre-discussion, personal choices of its members.

Reasons for this phenomenon are numerous; however, three distinct possibilities are often mentioned in the literature. Firstly, the phenomenon may occur because discussion with like-minded others increases certainty. Secondly, some individuals may moderate their views in anticipation of opposing viewpoints (see Walker & Main, 1973) and, thirdly, the "reciprocal influences of members on one another reinforce and strengthen extant orientations" (Feldman & Newcomb, 1969, p. 223).

Although each of these approaches has utility in some circumstances, each also has significant limitations for the educational decision-maker concerned with programme improvements. The Delphi is a technique for collecting judgements that attempts to

overcome the weaknesses implicit in relying on a single expert, a one-shot group average, or round-table discussion.

Irrespective of the focus area, the question of how critical decisions are made must impact directly on the results of the decision-making process, the decisions or ideas themselves. Subsequently, if the process is poorly conceived, it could be argued that the results may be inconsequential and invalid, thus failing to meet the assumptions behind the decision-making process and ultimately the needs of the client (Clayton, 1992).

Decision-making is a necessary task of all management functions. Critical decisions—that is, decisions which can either positively or negatively affect the overall functioning of organisations—demand a level of human endeavour and intellectualising which go above and beyond those decision-making activities of a daily or routine nature. Research supports the contention that credible and reliable results can be obtained through professional judgement and group decision-making processes (Reeves & Jauch, 1978; Zane *et al.*, 1982; Uhl, 1983, Babbie, 1990).

Moore (1987 pp. 15–17), provides four reasons why using a group of people rather than an individual is more desirable in conducting applied social research:

- (1) It is logical that if you properly combine the judgment of a large number of people, you have a better chance of getting closer to the truth.
- (2) It is desirable to use groups in order to understand social phenomena by obtaining the views of the actors.
- (3) It is often beneficial to use groups if you are concerned about the consequences of your research. If your goal is to solve a problem of a particular group, it is reasonable to believe that the group is more likely to accept your advice (or research findings) if they have participated in the research process.
- (4) Complex, ill-defined problems often can be addressed only by pooled intelligence.

Reliable information about current and future events and needs is necessary in corporate planning, in defence and general government planning and in institutions such as education (Linstone & Turoff, 1975). The decision process on issues as large as teacher training, for example, is generally left to administrators, policy-makers and educational theorists. It is generally assumed that groups such as these have formed a degree of vision and insight not found common to most people, and that their pooled wisdom is a formidable knowledge base on which to predict future trends, and that the underlying construct to this assumption involves some form of a group decision-making process (Uhl, 1983).

Typically, there are three group decision-making processes used for creative or judgemental problem solving: Nominal Group Technique (NGT), Interacting Group Method (IGM) and the Delphi technique (Delbecq *et al.*, 1975). NGT provides a forum for a group of people to develop and write ideas in a face-to-face situation, but the development of ideas is strictly individual and independent of other group members. Using a monitoring team, NGT pools and collates the ideas of the individuals and usually presents them to the group as a whole in writing often using a blackboard or flipchart. Ranking or rating then occurs and a group decision is made based upon a statistical criterion for aggregating the individual judgements.

IGM is a process in which the individuals openly discuss their ideas with each other, give feedback and analyse one another's work. Simply stated, IGM is a traditional

brainstorming exercise. An endpoint is reached when the group arrives at a level of agreement.

The Delphi technique is similar in nature to NGT, but has characteristics not found in either NGT or IGM. Firstly, idea generation by individuals is not only individual and independent, but also isolated and anonymous. Secondly, communication between individuals is orchestrated by a director and occurs via written questionnaires and feedback reports. Group decision-making exercises also pose a variety of logistic concerns such as the ability to bring together a large group of people often required for both NGT and IGM methods.

Delphi provides a communication medium whereby individuals can participate without needing to travel, often long distances, to a group meeting place. Further, as mentioned previously, individuals participate anonymously as a strict requirement of the Delphi process. Anonymity substantially reduces the social-emotional behaviour often found when using other methods, which allows participants to focus on task-oriented activities.

Uhl (1983) maintains that there are several problems commonly resulting from face-to-face discussion which can jeopardise individual objectivity and the integrity of an individual's response. Firstly, while group opinion is highly influenced by dominant individuals who usually monopolise a discussion, there is little correlation between verbosity and knowledge of the subject matter under consideration. Secondly, group discussion while appearing to be problem-oriented is often irrelevant or biased, as it is usually more concerned with individual or group interests than with problem solving. Thirdly, individual judgement can be distracted by group pressure to conform, an effect referred to as acquiescence whereby individuals may agree to an otherwise unacceptable position or idea without protest.

The Delphi method tries to obtain the most reliable consensus of opinion of a group through a series of intensive questionnaires interspersed with controlled feedback. The technique involves repeated questioning of the individuals and avoids direct confrontation of group members with each other.

For the purpose of clarification, Linstone and Turoff (1975, p. 3) capture the previously mentioned ideas in their definition:

Delphi may be characterized as a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem.

### **Delphi Background**

The derivation of the term 'Delphi' relates to the 'Delphic Oracle', an ancient Greek myth which held that a 'chosen one' on the island of Delphi was able to predict the future with infallible authority. Delphi was originally used to forecast technological developments; thus, like the oracle, it was used to look into the future.

The use of Delphi evolved from experimental research conducted by Dalkey and Helmer (1963) when working for the RAND Corporation. Called Project Delphi, the initial experiment 'was designed to apply expert opinion to the selection, from the viewpoint of a Soviet strategic planner, of an optimal U.S. industrial target system and to the estimation of the number of atomic bombs required to reduce the munitions output by a prescribed amount' (p. 458).

## Applications

Subsequent research has seen the Delphi technique used to generate ideas and forecast changes for a wide variety of educational purposes, such as curriculum development (Reeves & Jauch, 1978; Hartman, 1981; Blair & Uhl, 1993; Volk, 1993; Wells, 1994; Kluttschkowski & Troth, 1995); the future of adult education (Leirman, 1996); nurse education (Hartley, 1995); institutional planning (Uhl, 1983); determining educational effectiveness (Roberts *et al.*, 1984); forecasting expectations relating to the condition emotional disturbance/behaviour disorder (Carpenter, 1985); forecasting effects of deinstitutionalisation and necessary educational services (Putnam & Bruininks, 1986); identifying conditions most likely to encourage full participation in non-formal education programmes (Spencer-Cooke, 1986); distance education (Miller & Husmann, 1994); assessing the goals of elementary school gifted programmes (Hickey, 1988); identifying features of effective in-service practices (Van Tulder *et al.*, 1988); identifying competencies (Clayton, 1992; Cannon *et al.*, 1992; Smith & Simpson, 1995; Thach & Murphy, 1995); vocational training (Hakim & Weinblatt, 1993); investigating future directions in education and inclusion for students with disabilities (Putnam *et al.*, 1995); teacher effectiveness (Stivers & McMorris, 1991); and marketing teacher education (Smith, 1992).

## Delphi Characteristics

According to Linstone and Turoff (1975), there are three types of Delphi: conventional, real-time and policy. In conventional Delphi, a team designs a questionnaire which is sent to a larger respondent group. After the questionnaire is returned, the monitor team summarises the results and, based upon the results, develops a new questionnaire for the respondent group.

The respondent group is usually given at least one opportunity to re-evaluate its original answers based upon examination of the group response. Real-time Delphi differs from conventional Delphi in that, rather than taking weeks to conduct the process, it occurs during the course of a meeting or conference. In policy Delphi, the decision-maker is not interested in having a group generate his decision, but, rather, in having an informed group present all the options and supporting evidence for his/her consideration. The policy Delphi is not a mechanism for making decisions as generating a consensus is not the prime objective.

## Selection of Expertise

The Delphi method requires that a panel of experts on the subject under study be selected. An expert is someone who possesses the knowledge and experience necessary to participate in a Delphi. "A nuclear physicist is an appropriate expert if the Delphi concerns atomic energy and a resident of a neighbourhood is an expert on what should be a community's goals" (Moore, 1987, p. 51). All experts may be included or a random or nonbiased sample of various types of expertise may be sought. Expertise, however, is the desired goal for panel selection and it is this feature which sets Delphi apart from other general forms of survey research. In some Delphis, 'representativeness' is the criterion for panel selection, in which case appropriate random sampling strategies must be used (e.g. Smith & Glass, 1987; Borg & Gall, 1989; Babbie, 1990).

Because Delphi is a tool to aid understanding or decision-making, it will only be an

effective process if those decision-makers who will ultimately act upon the results of the Delphi are actively involved throughout the process. Further to this, expert panel members must be sufficiently motivated to include the Delphi task in their schedule of daily competing tasks and are prepared to value the aggregation of judgements of the respondent panel to which they would not otherwise have access (Delbecq *et al.*, 1975). The critical issue in all of this is to identify the 'expert' qualifications of panel members.

The process of selecting experts is critical to the Delphi and serves to authorise the Delphi's superiority and validity over other less painstaking and rigorous survey procedures. Failure to adhere to carefully orchestrated selection procedures of experts could render the Delphi weak and thus make the product or decisions less attractive to clients or end users. Once the general characteristics of the desired panelists are agreed upon, the Delphi director needs to initiate a nomination process.

Nominations of well known and respected individuals from members within selected target groups should be solicited and, through a process of ranking and culling, highly ranked nominees become evident and form the basis for panel selection. The selection process itself can also be quite motivating as there is a degree of flattery associated with being nominated as an expert by one's peers.

### Panel Size

Depending on the purpose of the study, the complexity and the expertise required, the panel may be large or small and local, state, national, or international. Group size theory varies, but some general rules-of-thumb indicate 15–30 people for a homogeneous population—that is, experts coming from the same discipline (e.g. nuclear physicists)—and 5–10 people for a heterogeneous population, people with expertise on a particular topic but coming from different social/professional stratifications such as teachers, university academics and school principals (Delbecq *et al.*, 1975; Uhl, 1983; Moore, 1987).

### Distribution of Questionnaire

(See Appendix 1 for a general procedure.) This involves three phases. In Phase One, a stimulus such as a questionnaire is devised and mailed to each member selected soliciting their co-operation in the study as well as their opinions or answers to certain events or questions based on their experienced judgement (Uhl, 1983; Putnam & Bruininks, 1986; Moore, 1987). The responses from co-operating members are then translated into general generic statements about which consensus is then sought. These statements are further screened in an effort to eliminate duplicate statements and to prepare a combined listing of all statements. This synthesis of the responses into a new questionnaire concludes Phase One.

It has also been argued that questionnaires are not well suited for the study of human attitudes and behaviour because they can elicit unreliable and biased self-reports (Alexander & Becker, 1978). These authors suggest that judgements required of respondents are often too abstract and that respondents answer in terms of their own mental picture of the task. One solution to the problem is to make the stimulus presented to the respondent as concrete as possible (Nosanchuk, 1972; Neff, 1979). This stimulus condition would more closely approximate a real-life decision-making or judgement-making situation.

The application of a stimulus such as this is referred to as a vignette or scenario



(Nosanchuk, 1972; Neff, 1979). Using this approach, the Delphi director could provide panel members with a stimulus scenario and require them to generate ideas or statements from which the decision-making process could then commence. This approach, however, is far more time consuming and more demanding of the panel members, but the greater involvement, input and ownership of the ideas generated should result in a final product which substantiates use of such a strategy.

Phase Two consists of several steps:

*Step 1.* Each expert receives a copy of the list of statements or ideas and is asked to rate or evaluate each item by some such criterion as importance or probability of success (Delbecq *et al.*, 1975). This is best done by use of a five- or seven-point Likert scale using zero as a mid (neutral) value with both positive and negative values (see below). The use of a scale such as this allows the researcher to work within an interval or quasi-interval scale of measurement (see Smith & Glass, 1987; Huck & Cormier, 1996).

|            |          |        |       |         |
|------------|----------|--------|-------|---------|
| S Disagree | Disagree | Unsure | Agree | S Agree |
| - 2        | - 1      | 0      | + 1   | + 2     |

*Step 2.* The responses from Step 1 are then summarised, giving a measure of central tendency, for example, mean, median, or interquartile range (Delbecq *et al.*, 1975; Reeves & Jauch, 1978). Each respondent is then asked to reconsider their previous answer in light of the group's measure and to revise it if they so desire. If the new response lies outside the central tendency, respondents are asked to state their reasons (if they so choose) for thinking that the answer should be that much higher or that much lower than the majority judgement of the group (Cyphert & Gant, 1970; Clayton, 1992).

*Step 3.* As in the previous step, the responses are again summarised. In addition, a concise summary of the reasons presented in support of extreme positions is formulated. Respondents are again asked to revise their positions on the basis of the summary of responses and taking into consideration the arguments presented.

*Step 4.* If this step is necessitated, then the criticisms of the reasons previously offered are resubmitted to the respondents for any final revisions.

Phase Three requires the Delphi director to establish a criterion as to which statements or ideas provide a strong indication of expert consensus. From the commencement of the Delphi procedure, panelists are creating and/or identifying ideas they perceived to be important or critical to the focus area and, as such, the process of identifying followed by rating and re-rating is usually one of refining rather than defining the criterion. The measure of the refining process lies in the Likert scale value each panel member attributes to each idea. Within the possible range of values each respondent could give, a value of '+ 1' or higher, for example, may indicate that the panelist agrees to strongly agrees that a statement is critically important. The director may then set a criterion that only those statements which receive a mean rating of '+ 1' or above are indicative of expert and strong consensus.

Variations of the technique are possible. Phase One may be eliminated if consensus is desired on already determined objectives or for setting priorities. Steps Three and

Four may not be necessary if convergence or consensus has been reached at the end of Step Two. Success of this method depends upon the choice of experts, the sampling techniques used to obtain representative opinions and the way in which the technique is implemented (Campbell & Hitchin, 1968). Researchers intending to use a Delphi approach are encouraged to use a three or four round process to achieve maximum input and reflection from their panel members (Linstone & Turoff, 1975; Moore, 1987).

The Delphi has also had its share of critics. Sackman (1975), one of the most ardent critics of this methodology, raises several concerns related to the scientific nature of the Delphi process. Although Sackman (p. 33) poses several key questions relating to conventional Delphi, the research literature has signally failed to address Sackman's concerns. This paper presents an appropriate opportunity to do so. Each of these questions will be responded to in turn.

(1) *"Is the Delphi concept of the expert and its claim to represent valid expert opinion scientifically tenable, or is it overstated?"*

It is well acknowledged and socially validated that the concept of 'expert' or 'expertise' is both a social and scientific phenomenon. Expertise exists in various forms and, although it may be difficult to measure exactly, there are general characteristics of individuals who, in a given context, demonstrate a level of wisdom, insight, theory, practice, experience and analysis not found common to all individuals. It is these individuals to whom the term 'expert' is assigned. It is reasonable, therefore, to seek out individuals whose peers regard them in this light. In the final analysis, and assuming that appropriate selection processes have been followed, how the results are presented, discussed and analysed will determine whether the opinions or ideas generated by the respondents are overstated.

(2) *"Are Delphi claims of the superiority of group over individual opinion, and of the superiority of remote and private opinion over face-to-face encounter, meaningful and valid generalisations?"*

There is sufficient evidence revealing the substantial weakness of individual and face-to-face encounters (Moore, 1987; Forsyth, 1990) and more than sufficient evidence to accept that, at the very least, a group of experts would normally deliver not only a more superior product than one expert, but also a more generalised expert perception of the focus area.

(3) *"Is Delphi consensus authentic or specious consensus?"*

When consensus can be achieved by refinement through discussion and feedback rather than through alignment by acquiescence—an effect whereby individuals may agree to an unacceptable position without protest—its authenticity is far from specious and far more valid.

(4) *"Does Delphi anonymity reinforce scientific accountability or unaccountability in method and findings?"*

Anonymity is an important feature of the Delphi process, its intention being to disinhibit effective individual thought and analysis which may be influenced by persuasively stated opinions of significant others (Uhl, 1983). Anonymity is needed to support the contention that the collective wisdom of experts can be refined through controlled

feedback and discussion. Anything external which may unduly influence an individual's consideration only serves to weaken its validity.

(5) *"Does Delphi systematically encourage or discourage the adversary process and exploratory thinking?"*

Delphi seeks actively to encourage debate and exploratory thought without the face-to-face pressures of direct confrontation which may inhibit the consideration of novel ideas. Anonymity and geographical separation eliminate these pressures without weakening independent thought and the gradual formation of a considered opinion.

(6) *"Are Delphi questions, particularly forecasting questions, precise and meaningful?"*

Any question, which originates from the wisdom of hindsight and draws upon the cumulative wealth and experience of practice and research, seeking to make an inference about the future, must stand a greater probability of being answered more accurately than those questions which are ill thought-out and are poorly aimed, and which fail to account for current knowledge and relevant history.

(7) *"Are Delphi responses precise and ambiguous?"*

In some ways the very reason why Delphi is such a useful and potent methodology is that it systematically attempts to control for those elements which tend to militate against and weaken other forms of group decision-making methods. The precision and ambiguity of the Delphi response is ultimately an effect of the process which the Delphi director has used in much the same way as it is for any piece of research. That is, if you use the correct tools inappropriately, you may well end up with a poor or unreliable product.

(8) *"Are Delphi results meaningful and unambiguous?"*

The importance of the results is, again, an effect of the process itself. The potency of the results, however, must closely correlate with the selection of the panel members. High level, high calibre, expertise and judgement will be viewed and accepted by others from within the field if the end-users sense ownership and recognise the proposed authority of those who participated in the development of the recommendations.

(9) *"Is Delphi primarily concerned with collections of snap-judgement opinions of polled individuals from unknown samples, or is it concerned with coherent predictions, analyses, or forecasts of operationally defined and systematically studied behaviours or events?"*

The latter half of this question is Delphi's *sine qua non*.

(10) *"Does Delphi represent a critical tradition, or is it uncritically isolated from the mainstream of scientific questionnaire development and behavioural experimentation?"*

Evidence from the literature would indicate that Delphi is becoming more popular as a forecasting tool and as a reliable method for decision-making purposes, but its traditional use over the past 34 years may be equivocal. This paper has been written as one way of drawing the attention of researchers to this procedure so that it may become more academically mainstreamed. It is this author's contention that any equivocation regarding Delphi's use is more a result of unfamiliarity than with perceived weaknesses relating to its scientific substance.

(11) *"Does Delphi set a desirable or an undesirable precedent for interdisciplinary science in the professional planning and policy studies community?"*

Planning and policy development is the predicate of Delphi. Regardless of the nature or substance of inquiry or investigation, critical decision-making processes take place on a daily basis within all institutions. When individuals seek to evaluate, question, interpret or predict significant events, Delphi is a tool which would add value and potency to the decision-making process.

## Conclusions

The decision to use the Delphi technique should be based on the purpose or objective of a research study which wishes "to obtain the most reliable consensus of opinion of a group of experts" (Dalkey & Helmer, 1963, p. 458). If the objective is the identification of content based on expert consensus, then the Delphi technique is an appropriate choice as it may enhance the significant contributions of the panel.

As with any study involving the analysis of perception and expert opinion, a number of limitations can be apparent. The following limitations are considered to be particularly pertinent to a better understanding of Delphi:

- (1) The background and experiences of each panel member, which may have directly affected their decision-making, are generally beyond the control of a Delphi study.
- (2) Due to both personal and professional obligations, panel members are often limited in the amount of time each can dedicate to the decision-making process. This may effectively reduce each member's ability to consider and report on all dimensions under investigation.
- (3) The Delphi process aims to arrive at a level of consensus among the panel members. This consensus is assisted by the researcher providing feedback to each panel member of their previous rating together with a group measure of central tendency. The researcher may provide both individual ratings and the group mean rating. It is unknown whether the panel members think through their relative positions and work towards authentic consistency of opinion, or if they are effectively pressured into conformity. The issue of individual acquiescence or tacit approval of the group's perception cannot be measured, and although responses in each of the rounds may be reliable, they may not constitute valid personal responses.
- (4) Although the results from a group of experts can yield substantive and thought-provoking results, they may not be an exhaustive nor all-inclusive set of ideas. A study may yield a valuable source of information; however, the value of the information is for the individual reader to decide and is limited due to the constraints imposed by the panel selection, as well as by the backgrounds, experiences and biases of each member.
- (5) It is possible that the researcher's analysis of the results is not the only interpretation which can be made and that some distortion may occur due to the researcher's own biases.

Even in the face of limitations, and researchers are cautioned to take into account these limitations when considering the results of Delphi studies, the Delphi has great strength and utility. It collects and organises judgements in a systematic fashion. It gains input, establishes priorities and builds consensus. It organises and helps to focus dissent,

turning this group effect into a window of opportunity. In short, Delphi cannot be overlooked as a useful and potent tool when attempting to harness expert opinion for critical decision-making tasks in education.

*Correspondence:* Mark J. Clayton, Macquarie University, Special Education Centre, Macquarie University, Sydney, New South Wales 2109, Australia.

## REFERENCES

- ALEXANDER, C.S. & BECKER, H.J. (1978) The use of vignettes in survey research, *Public Opinion Quarterly*, 42, pp. 93–104.
- BABBIE, E. (1990) *Survey Research Methods*, 2nd Edn (California, Wadsworth)
- BERDIE, D.R., ANDERSON, J.F. & NIEBUHR, M.A. (1986) *Questionnaires: design and use*, 2nd Edn (New Jersey, The Scarecrow Press).
- BLAIR, S. & UHL, N.P. (1993) Using the Delphi method to improve the curriculum, *Canadian Journal of Higher Education*, 23, pp. 107–128.
- BORG, W.R. & GALL, M.D. (1989) *Educational Research: an introduction*, 5th Edn (New York, Longman).
- BRANDSTATTER, H., DAVIS, J.H. & STOCKER-KREICHGAUER, G. (1982) *Group Decision Making* (London, Academic Press).
- CAMPBELL, R.M. & HITCHIN, D. (1968) The Delphi technique: implementation in the corporate environment, *Management Services*, pp. 37–42.
- CARPENTER, R.L. (1985) Emotional disturbance/behavioral disorders and a delphic approach to the future, *Behavioral Disorders*, May, pp. 198–210.
- CLAYTON, M. (1992) Identifying critical entry competencies for beginning special educators: the Delphi study, *Unpublished doctoral dissertation*, University of Illinois, Champaign-Urbana.
- CYPHERT, F.R. & GANT, W.L. (1970) *The Delphi Technique: A tool for collecting opinions in teacher education*, Paper presented at the AERA Symposium, Exploring the potential of the Delphi technique by analyzing its application, Minneapolis, MN.
- DALKEY, N. & HELMER, O. (1963) An experimental application of the Delphi method to the use of experts, *Management Science*, 9, pp. 458–467.
- DELBECQ, A.L., VAN DE VEN, A.H. & GUSTAFSON, H. (1975) *Group Techniques for Program Planning* (Glenview, IL, Scott, Foresman and Company).
- FELDMAN, K.A. & NEWCOMB, T. (1969) *The Impact of College on Students* (San Francisco, CA, Jossey-Bass).
- FORSYTH, D.R. (1990) *Group Dynamics*, 2nd Edn (California, Brooks/Cole).
- HAKIM, S. & WEINBLATT, J. (1993) The Delphi process as a tool for decision-making: the case of vocational training of people with handicaps, *Evaluation and Program Planning*, 16, pp. 25–38.
- HARTLEY, M.E. (1995) The development of module evaluation: a delphi approach, *Nurse Education Today*, 15, pp. 267–273.
- HARTMAN, A. (1981) Reaching consensus using the delphi Technique, *Educational Leadership*, March, pp. 495–497.
- HICKEY, G. (1988) Goals for gifted programs: perceptions of interested groups, *Gifted Child Quarterly*, 32, pp. 231–233.
- HUCK, S.W. & CORMIER, W.H. (1996) *Reading Statistics and Research*, 2nd Edn (New York, Harper Collins).
- KLUTSCHKOWSKI, F. & TROTH, W.A. (1995) Components for a substance abuse counselor education curriculum: a Delphi study, *Journal of Addictions and Offender Counseling*, 15, pp. 34–41.
- LEIRMAN, W. (1996) Euro Delphi: a comparative study on the future of adult education in 14 countries between 1993 and 1995 *International Journal of Lifelong Education*, 15, pp. 125–138.
- LINSTONE, H.A. & TUROFF, M.T. (1975) *The Delphi Method: techniques and applications* (London, Addison-Wesley).
- MILLER, M.T. & HUSMANN, D.E. (1994) Strategies for improving instructional delivery in distance education programs, *Journal of Adult Education*, 22, pp. 23–29.
- MOORE, C.M. (1987) *Group Techniques for Idea Building* (California, Sage).

- NEFF, J.A. (1979) Interactional versus hypothetical others: the use of vignettes in attitude research, *Sociology and Social Research*, 64, pp. 105–125.
- NOSANCHUK, T.A. (1972) The vignette as an experimental approach to the study of social status: an exploratory study, *Social Science Research*, 1, pp. 107–120.
- PUTNAM, J.W. & BRUNINKS, R.H. (1986) Future directions in deinstitutionalization and education: an Delphi investigation, *Exceptional Children*, 53, pp. 55–62.
- PUTNAM, J., SPIEGEL, A. & BRUNINKS, R. (1995) Future directions in education and inclusion of students with disabilities: a Delphi investigation, *Exceptional Children*, 61, pp. 553–576.
- RASP, A. (1973) Delphi: a decision-maker's dream, *Nation's Schools*, 92, p. 1.
- REEVES, G. & JAUCH, L.R. (1978) Curriculum development through Delphi, *Research in Higher Education*, 8, pp. 157–168.
- ROBERTS, M.D., GRIMES, R.M. MOSELEY, S.K. & BRUHN, J.G. (1984). Development of indices of effectiveness: a quantification of the accrediting process, *Journal of Allied Health*, February, pp. 13–21.
- SACKMAN, H. (1975) *Delphi Critique: expert opinions, forecasting, and group process* (Lexington, M.A., D.C. Heath).
- SMITH, C.L. (1992) National issues and concerns in marketing teacher education, *Marketing Educators' Journal*, 30, pp. 75–89.
- SMITH, K.S. & SIMPSON, R.D. (1995) Validating teacher competencies for faculty members in higher education: a national study using the Delphi method *Innovative Higher Education*, 19, pp. 223–234.
- SMITH, M.L. & GLASS, G.V. (1987) *Research and Evaluation in Education and the Social Sciences* (New Jersey, Prentice-Hall).
- SPENCER-COOKE, B. (1986) Conditions of participation in rural, non-formal education programmes: a Delphi study, *Educational Technology: Focus on People*, 26, pp. 115–124.
- STIVERS, J. & MCMORRIS, R.F. (1991) Relating a test for teachers to research literature on teaching effectiveness, *Journal of Personnel Evaluation in Education*, 5, pp. 31–53.
- THACH, E.C. & MURPHY, K.L. (1995) Competencies for distance education professionals, *Educational Technology Research and Development*, 43, pp. 57–79.
- UHL, N.P. (1983) Using the Delphi technique in institutional planning. Using research for strategic planning, *New Directions for Institutional Research*, No. 37 (San Francisco, Jossey Bass).
- VAN TULDER, M., VEENMAN, S. & SIEBEN, J. (1988) Features of effective in-service activities: results of a Delphi study, *Educational Studies*, 14, 209–223.
- VOLK, K. (1993) Curriculum development using the Delphi technique, *Technology Teacher*, 52, pp. 35–36.
- WALKER, T.G. & MAIN, E.C. (1973) Choice shifts in political decision-making: federal judges and civil liberties cases, *Journal of Applied Social Psychology*, 2, pp. 39–48.
- WELLS, J.G. (1994) Establishing a taxonomic structure for the study of biotechnology in secondary school technology education, *Journal of Technology Education*, 6, p. 1.
- ZANE, T., SULZER-AZAROFF, B., HANDEN, B.L. & FOX, C.J. (1982) Validation of a competency-based training program in developmental disabilities, *The Journal of the Association for the Severely Handicapped*, 7, pp. 21–31.

## Appendix 1: Generalised Research Procedure

### Round One

Prior to mailing Round One, a letter is sent to each member informing them that the first round would be mailed on or before a set date. In addition, the letter serves to stimulate the participants and requests that each take a personal interest in the study. The first round survey is mailed to each participant. The Round One survey consists of a number of sections, each section serving as a task to be completed by each member. The first section contains the purpose and rationale; the second section contains the directions; the third section contains the instructions required to complete the Delphi effectively and efficiently.

A cover letter is attached to the first round survey that welcomes participants to the study and explains the general procedures to be followed. The cover page should be developed using graphics software to ensure a pleasant, artistic and professional presentation (Berdie *et al.*, 1986, Babbie, 1990). Each participant is asked to read the rationale, directions, instructions etc., before attempting to complete their survey.

At least 3 weeks after mailing the first round, each member is telephoned to ensure that they received

the package and to prompt their completion and return of it. This should be followed up on a two weekly basis until all responses have been received. Fax messages and email are used where available.

Once received, the responses identified in Round One from each panel are typed into a word-processor and printed. The final product is then evaluated by the principal researcher and a group if necessary to determine redundancies and identify opinions that could be collapsed. Opinions perceived to have these characteristics, and which are agreed to by all reviewers, are then eliminated from inclusion in Round Two.

### Round Two

The Round Two survey is mailed on a date deemed appropriate following the analysis of the first round and usually consists of three sections. Section One contains a set of instructions. Section two consists of the complete set of opinions etc. identified in Round One minus those eliminated by the reviewers and a Likert scale placed immediately below each statement etc.

An explanation for the Likert scale is provided in the instruction page indicating the values. Panel members are asked to rate each statement by checking their perceived agreement with each statement. The scale can have as the lower anchor, strongly disagree (SD) - 2; and the higher anchor, strongly agree (SA) + 2. Neutral (N) is assigned a zero value, and disagree (D) and agree (A) - 1 and + 1, respectively. The zero point provides a neutral option to panel members and is intended to overcome forced-choice responding and provide members every possible option in rating (Berdie *et al.*, Anderson, & Niebulur, 1986).

Section Three can provide space in which members are able to write additional statements if they so desire, but they are not compelled to do this. Panel members are asked to mail the second round at least 5 days after receiving it.

The data obtained from Round Two is usually analysed using descriptive statistics. A criterion can be set such that a statement scoring a mean of '1' or higher, for example, is set as the criterion for including the statements/opinions in the third round. The statements not achieving the criterion are eliminated. The reason for setting such a liberal criterion is to ensure that panel members make the decisions and not the researcher, thereby reducing any vested interest the researcher may have (e.g. Uhl, 1983).

### Round Three

The Round Three instrument is mailed on a date determined to be appropriate following the analysis of the second round and usually consists of three sections. The first section contains the instructions which includes an explanation of the symbols relating to the Round Two ratings. The second section contains the statements/opinions remaining from Round Two and contains all the elements from Round Two plus two new components. Placed above the respective Likert scale is the group mean rating for each competency indicated by an (X). This shows the relative position of the group mean in relation to the set of boxes and is considered easier to understand than providing a numerical value such as 1.325. In addition to the symbol for the group mean, a red (V) can indicate where each individual member rated that statement/opinion in Round Two.

|    |   |   |   |    |  |
|----|---|---|---|----|--|
|    |   | V |   | X  |  |
| SD | D | N | A | SA |  |

The purpose of providing these two scores is to indicate to each member where they lie in relation to the groups' perception. Each member is then given an opportunity to re-rate each statement with knowledge of the groups' decision. It is explained to each member that although consensus is desirable, they should not feel compelled to rate according to the groups' rating. However, members can be advised that if they differ markedly to the mean rating they should give careful reappraisal to that statement. In addition, members can be given the opportunity to explain their reasons for their ratings on the third round, but are not compelled to do this (Cyphert & Gant, 1970).

The third section of the third round asks each panel member for general demographic information. Members are asked to mail the third round 2 days after receiving it. Usually included in the third round package of materials is a copy of the complete set of statements minus the Likert scales. Panelists are asked to keep this copy on hand as any discrepancies emerging from the third round can be dealt with by telephone, fax or email to expedite the communication process.

*Communication Process*

During the course of the study, numerous telephone calls, fax and email messages are received and sent. Numerous telephone calls can be made both at the onset of the study and during its course. Both home telephone numbers and work numbers are required. The importance of maintaining close contact with participants in a study of this nature cannot be stressed too highly, and the high response rates achieved by Delphi studies are likely to have resulted from maintaining close, cordial and frequent contact.