

USING QUESTIONNAIRES TO ASSESS MOTIVATION IN SECOND LANGUAGE CLASSROOMS

PATRICIA O'BRYEN

University of Hawai'i

Questionnaires for measuring motivation and attitudes in second language (SL) learning can be adapted from existing instruments or they can be constructed for use in other contexts. The purpose of this review is to clarify the methodology for carrying out SL motivation research with questionnaires in classrooms. Issues of validity raised in past research are summarized. Twelve questionnaires used in English as a second language (ESL) and foreign language learning contexts (Clement & Kruidenier, 1983; Clement, Smythe, & Gardner, 1976; Dornyei, 1990; Ely, 1986a, 1986b; Gardner & Smythe, 1981; Gliksmann, Gardner, & Smythe, 1982; Labrie & Clement, 1986; Pennington & Yue, 1994; Pierson, Fu, & Lee, 1981; Roger, Bull, & Fletcher, 1981; Samimy, & Tabuse, 1992) are considered for potential use in classroom studies. Depending on whether the research is exploratory or confirmatory, factor analysis, LISREL (linear structural relationships) modeling, or MDS (multidimensional scaling) statistical procedures may be used to analyze the multivariate response data, and the results will take different configurations. Perceptions gained from the review guided the author in proposing eleven steps for developing a motivation questionnaire.

INTRODUCTION

Teachers of second languages almost all agree that successful language learners are highly motivated, but few have measured the motivations of their students systematically. Until recently, only researchers have assessed motivation and the other affective variables that go with second language (SL) learning. Observations of individual learners are valuable, but most of the data come from cross-sectional studies conducted in classroom settings. Teachers, moved by an interest in facilitating instruction and in having learners persevere in their study of languages, may also want to measure the motivations of their students.

Measurement links theoretical rationalizations and empirical evidence derived from observational data, whether the data are coded qualitatively or quantitatively (Messick, 1989). Researchers typically measure motivation in SL learning with questionnaires

designed for specific learners. Existing questionnaires can sometimes be used in other studies, or they may be suitable in another context if they can be modified. New questionnaires are constructed if none of the existing instruments are appropriate for the learners of interest, or if changes in theories cause research efforts to be directed towards other aspects of motivation. Thus, experience and skill in working with questionnaires are key issues when measuring SL motivation.

Administering a questionnaire does not require large amounts of skill or time, but motivation can only be measured indirectly, and the processes leading to and after the use of even a simple questionnaire can be difficult and time-consuming (Nunan, 1992). This review looks at twelve studies from the literature on SL motivation research, in which data were collected with questionnaires, with the goal of clarifying the process of using questionnaires to assess motivation in SL classrooms. Specifically, the first section of the review gives a short introduction to research on the role of motivation and attitude in SL proficiency. Then, some background on using questionnaires to assess motivation precedes a summary of a debate in past research on self-report measures used to assess motivation in SL learning. The next section brings out differences between the questionnaires used in research, including those from the selected studies, comparing types and numbers of learners, the main constructs measured, and the rating scales used for the responses. This is followed by short introductions to some of the statistical analyses used with questionnaire data, factor analysis in particular. The review concludes with steps for developing questionnaires to use in SL classroom contexts.

Motivation, Attitude, and Proficiency in SL Research

Theoretical rationale is used inductively in research, to understand specific observations in terms of generalizations, or deductively, to regularize observations and make predictions to arrive at interim explanations of a process (Larsen-Freeman & Long, 1991; Messick, 1989). In SL research, motivation and attitude, affective factors associated with second language learning, help to explain observations of the relative degrees of proficiency achieved by individual SL learners (Larsen-Freeman & Long, 1991). For example, individuals who want to learn another language and who devote

substantial effort to doing so are generally considered to be motivated, and they are often observed to be more proficient than SL learners who do not have these characteristics, other factors (such as age) being equal.

Researchers do not yet agree about whether or not a set of items can be established that define the motivation construct, but evidence based on the responses to questionnaires suggests that motivation in SL learning can be described in terms of a large number of indirectly observed variables that can be reduced to several underlying factors (Hatch & Lazaraton, 1991). Distinctions between integrative and instrumental orientations, intrinsic and extrinsic motivation, and other explanations of empirical evidence have supplemented considerations of parsimony and elegance in the development of SL motivation theories (see, for example, Brown, 1990; Clement, Dornyei, & Noels, 1994; Ramage, 1990; Skehan, 1989).

Although theories of motivation are still evolving, there are a number of reasons for linking attitude to motivation in language learning research. Attitudes develop in an individual over time and can be measured. They are associated with interests, motives, aesthetic appreciation, values, goals, ideals, character, morale, and social distance, characteristics of individuals that are hypothesized to affect SL learning (Ames & Archer, 1988; Duda & Nicholls, 1992; Evans & Jarvis, 1986; Tremblay & Gardner, in press). In addition, the large number of working definitions associated with attitude allow great flexibility in developing the items on questionnaires that serve as the independent variables in many of the studies.

Some SL researchers have treated attitudes and motivation separately (see Larsen-Freeman & Long, 1991), but motivation and attitudes to the learning situation (i.e., "Integrativeness") are central ideas of the socio-educational model, which proposes causal relationships between attitudes and motivation and between motivation, aptitude, and achievement in SL learning. Widely researched by Gardner and his associates, the model has contributed notably to current understanding of the role of motivation in SL learning in formal contexts (for a review, see Crookes & Schmidt, 1991; Gardner, 1980; Skehan, 1989). For example, the relationships proposed in the socio-educational model have been

confirmed in a LISREL analysis of questionnaire data from two different SL settings (Gardner & Tremblay, 1994a; Kraemer, 1993; Tremblay & Gardner, in press).

Proficiency scores in the language being studied are often the dependent variables in motivation research, but proficiency must also be measured indirectly. Views about what constitutes SL proficiency differ in testing theory (Bachman, 1990; Chapelle & Douglas, 1993), and this is reflected in SL motivation studies. For example, in past research, language proficiency has been described as "the actual acquisition of skill in the target language", operationalized as scores on a cloze passage (Oller, 1981a, p. 228), and as "achievement", operationalized as grades in the target language, in other research conducted at about the same time (Gardner, 1980). In more recent motivation studies, notions of proficiency have also included grades on an essay (Tremblay & Gardner, in press) and SL proficiency from the learner's point of view (Dornyei, 1990).

Since motivation, attitude, and SL proficiency are underlying constructs and not directly observable, researchers interpret observed consistencies in these characteristics in terms of assigned scores, whether the evidence is based on observational data that are scored qualitatively or quantitatively (Messick, 1989). For example, indirect measures of motivation, such as nomination of playmates, have been used in studies with preliterate learners (Strong, 1984). Inferences about motivation in SL learning in formal contexts have been largely based on the scored responses to other indirect measures, self-reports and questionnaires.

Questionnaires in SL Classroom Motivation Research

Although the motivation of learners has been assessed qualitatively in a number of case studies (Csikszentmihalyi & Nakamura, 1989; Schumann, 1978), statistical studies with questionnaires have been used more widely to aid in the operationalization of motivation and to assess motivation in second and foreign language classroom contexts. A questionnaire is operationally defined as a research instrument, a collection of independent variables, generally handed to the respondents and filled in by them without any help from the researcher (Bailey, 1987). In this sense, it differs from the interview schedule,

sometimes also referred to as a questionnaire, which is filled in by an interviewer as the respondent answers spoken questions.

Variables on questionnaires are often organized as one or more sets of items (scales) that appear to be closely related to one aspect of a construct. Generally, the larger the number of items used to assess the construct, the more reliable and valid the inferences based on the questionnaire scores are likely to be (Messick, 1989). Oppenheim's (1966) observation that attitudes and motivation do not necessarily exist along linear continua, but that it is convenient to think of them in this way for measurement purposes, is still relevant today. Likert scales (strongly disagree, moderately disagree, slightly disagree, neutral, slightly agree, etc.) are usually used as response formats for assertions related to an attitude or motivation, but multiple-choice formats (Gardner, Clement, Smythe, & Smythe, 1985), bipolar adjective scales (Oller, 1977), and semantic differential scales (Clement, Smythe, & Gardner, 1976) have also been used to quantify variables on questionnaires.

The Attitude/ Motivation Test Battery (AMTB) is perhaps the best known motivation questionnaire used in a SL learning classroom context (Gardner, et al., 1985). Originally designed for use with Anglophone Canadian high school learners of French as a second language, the AMTB includes assertions about attitudes towards speakers of French and toward learning French, as well as self-reports of efforts expended in learning the language. Scores on motivation, attitude, and aptitude scales are the independent variables, and the measure of achievement, the dependent variable, is usually grades in French. Numerous administrations of the AMTB have led to the formulation of the socio-educational model and its hypothesized relationships.

Validation and Measures of Affective Factors

Using self-report measures such as questionnaires to assess affective variables in language learning has not been without controversy. Since theories of motivation are still evolving, and since evidence is always incomplete (Messick, 1989), it is perhaps not surprising that inferences based on theoretical rationales and empirical evidence are subject to error. In past research, Oller and his associates found negative correlations between

items reflecting positive attitudes and SL proficiency among learners of English in foreign and second language learning contexts (Oller, Baca, & Vigil, 1977; Oller, Hudson, & Liu, 1977).

Oller and Perkins (1978a, 1978b; see also, Oller, 1981a, 1981b, 1982) interpreted these results as evidence that the shared variance between affective data and language proficiency scores was due to intelligence, an underlying general factor of language proficiency. That is, if the questionnaires were in the target language, they might be rather weak measures of SL proficiency. If they were written in the L1, they might be weak measures of intelligence, because three traits could be operating to produce construct-irrelevant variance: an approval motive, self-flattery, and response set. The approval motive causes variance when learners give the responses they think are the expected ones, instead of those which reflect their true attitudes. Self-flattery is based on what learners believe are desirable and undesirable characteristics, and is posited to cause variance when bipolar adjective-type scales are used. Response set, when learners try to be consistent in their responses, may account for some of the common variance in responses to related items on a questionnaire. Thus, when self-reported attitude data were used to predict language proficiency, it was speculated, the extraneous variance would produce spurious relationships (Oller & Perkins, 1978a).

The authors insisted that self-report measures contained construct-irrelevant variance, which seriously threatened their internal validity (Oller & Perkins, 1978a, 1978b; see also, Oller, 1981a, 1981b, 1982). This was not a minor dissent because they were led to a position almost diametrically opposed to Gardner (1980; Gardner & Glikson, 1982) with respect to the independence of affective variables and SL proficiency. For Gardner, affective variables (AMI scores, a composite of the attitude and motivation scale scores on the AMTB) and aptitude variables (Modern Language Aptitude Test, or MLAT, scores) were better predictors of the criterion (grades in French) than either set of scores alone. For Oller, affective variables (a combination of attitude statements and bipolar adjective scales) and the criterion variable (a cloze passage) seemed to be measuring the same thing (i.e., verbal intelligence). As a consequence, the primary question for Oller became the validity of the self-report type measures of affect.

Gardner (1980) defended the use of the AMTB, which is not written in the target language and is a different type of self-rating than the scales used by Oller and his associates. He pointed out that the ability to predict SL achievement from affective variables was a generalization, so inconsistencies reported in other studies did not invalidate the interpretation that attitude and SL proficiency were related. In addition to sociocultural differences among the learners in the studies, Gardner offered other reasonable explanations of the negative findings in Oller et al.'s empirical studies, which included factors affecting the power of an analysis (Lazaraton, 1991), that is, the possibility of making Type I or Type II errors when making inferences from the data, small sample sizes, and the small number of variables used to assess attitude.

To determine the degree to which two measures are spuriously correlated, two conditions must be met: the traits must be differentially present in different learners, and they should affect total scores or ratings on the two measures to a comparable degree (Upshur, Acton, Arthur, & Guiora, 1978). In other words, discriminant and convergent evidence need to be given to support the validity of inferences made from sets of correlated scores. Gardner (1980) provided a detailed explanation of how he estimated proportions of shared variance among the scores on the AMTB and grades in French, obtained from thousands of learners of French as a second language in Canadian classroom contexts, along with an extensive discussion of how findings provided evidence on content validity, predictive validity, and convergent and discriminant construct validity.

To provide evidence in support of content validity, Gardner and Gliksman (1982) checked the representativeness of the attitude and motivation variables on the AMTB by calculating estimates of internal consistency of the items. Predictive validity evidence was provided by interpreting strength of association coefficients for the results of correlation studies, in which AMI scores were shown to account for approximately 14% of the variance in grades in French, and MLAT scores were shown to account for approximately 17% of the variance in the grades. Using partial correlation procedures, Gardner showed that the AMI and MLAT scores were relatively independent, but that the scores enhanced prediction of the criterion variable when they were combined. Gardner

and Gliksman (1982) interpreted these findings to conclude that the measures of attitude and motivation on the AMTB were substantially related to SL achievement.

For convergent construct validity, the authors cite findings that scores on the AMTB motivation scale were consistently highly correlated with aural comprehension scores, fluency and pronunciation scores, and grades in French. For discriminant construct validity, it had been important to show the independence of motivation/attitudes from aptitude, otherwise it could be said that the affective factors were important **because** they covaried with the ability to learn languages. Factor analysis confirmed that attitude/motivation and aptitude were independent factors. The relationship between affective variables and grades in French did not generalize to grades in other courses.

Gardner and Gliksman's (1982) analysis had superior power because they had a much larger sample size, they had assessed the construct with a larger number of items, and by reporting strength of association coefficients, they had provided an indication of the importance of the relationships (Lazaraton, 1991). As a result, they concluded that Oller and Perkins' hypotheses were not supported by evidence from research with the AMTB (Gardner & Gliksman, 1982).

Threats to validity have been discussed by Messick (1989) and by Cronbach (1988), among others. These authors also invoke the concepts of convergent and divergent evidence to support validity judgments. Validity is defined as "an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment" (Messick, 1989, p. 13). This formulation suggests that validity is a property of scores, not of measures. Although a large amount of the variance in questionnaire scores has yet to be identified, the need to establish the validity of the questionnaires used to assess affective variables, stressed by Oller, is not a genuine concern.

One of the outcomes of the debate was a move by SL researchers toward factor analysis and other advanced statistical procedures to attempt to increase the power of their analyses and to be able to account for larger amounts of the unexplained variance obtained using data from questionnaires, and to include other traits hypothesized to influence

successful language learning in motivation studies. The debate may also have been influential in broadening the scope of empirical studies to include more foreign and second language learning contexts.

A Sampling of SL Motivation Questionnaires

Looking at the way in which motivation has been researched in other contexts is useful. Occasionally, SL motivation studies, with their questionnaires, are published in the research literature. The use of one of these ready-made instruments is potentially the simplest method of researching motivation with other SL learners. Twelve motivation studies using questionnaires in SL learning contexts were selected from the literature as examples (Clement & Kruidenier, 1983; Clement et al., 1976; Dornyei, 1990; Ely, 1986a, 1986b; Gardner & Smythe, 1981; Gliksmann et al., 1982; Labrie & Clement, 1986; Pennington & Yue, 1994; Pierson et al., 1981; Roger et al., 1981; Samimy & Tabuse, 1992). Details about the subjects, languages and principal constructs for each of the questionnaires are presented in Table 1. A comprehensive overview of the 12 studies is given in the Appendix.

Table 1
Language Learning Motivation Questionnaire Studies

Study	Description of respondents	Main constructs measured
1. Clement, Smythe, & Gardner 1976	153 Grade 10 and 151 Grade 11 Francophone ESL students: Montreal	Attitudes towards English speakers, interest in foreign languages, integrative and instrumental orientation, motivational intensity
2. Pierson, Fu, & Lee 1980	466 Grade 10 Chinese EFL students: Hong Kong	Attitudes and achievement
3. Roger, Bull, & Fletcher 1981	86 students of French, 13-14 years old: York, UK (pilot study)	Attitudes and motivation

- | | | |
|--|---|---|
| | 938 students of French,
13-14 years old:
Leeds, and York, UK
(main study) | |
| 4. Gardner & Smythe
1981 | 552 students of French
SL, Grades 7-11:
Canada
(pilot study) | Attitudes, motivation,
achievement in French |
| | 1,521 students of French
SL, Grades 7-11:
Canada
(main study) | |
| 5. Gliksmann,
Gardner, & Smythe
1982 | 149 students of French,
Grades 9, 10, 11:
Ontario, Canada | Attitudes, integrativeness,
motivation, desire to learn
French, achievement in
French, participation in the
classroom |
| 6. Clement &
Kruidenier
1983 | 871 Grade 11 Anglophone
and Francophone
students of French,
English, and Spanish:
Quebec, Ottawa, and
London, Canada | Reasons for studying a
second language,
integrative orientation,
instrumental orientation |
| 7. Ely
1986a | 50 University students of
Spanish: California
(pilot study) | Language learning
motivation |
| | 75 University students of
Spanish: California
(main study) | |
| 8. Ely
1986b | 50 University students of
Spanish: California
(pilot study) | Language class discomfort,
strength of motivation, risk
taking and sociability |
| | 75 University students of
Spanish: California
(main study) | |

9. Labrie & Clement 1986	95	Grade 9 Francophone ESL students: New Brunswick, Canada	Ethnolinguistic vitality, attitudes, motivation, self-confidence, inter- ethnic contact
10. Dornyei 1990	34	Hungarian adult EFL students: Hungary	Language use, intentions, beliefs, values, interests, attitudes
11. Samimy & Tabuse 1992	68	University students of Japanese: Midwest USA	Risk taking, sociability, discomfort, motivation, attitude, grades
12. Pennington & Yue 1994	285	Chinese EFL students, Grades 7-12: Hong Kong	Attitudes and achievement

Note. ESL = English as a second language; EFL = English as a foreign language

Similarities and Differences

Respondents. Generally, the respondents were junior high and high school students, although a few were university freshmen or adult ESL learners. Almost all of the questionnaires had been given to large numbers of students (ranging from 50 to more than 1500) in single administrations. Only one of the studies was conducted longitudinally (Samimy & Tabuse, 1992). Longitudinal studies of motivation in language learners require considerably more time and planning than single administrations, because the learners are usually in school settings which follow quarter or semester terms which will affect the scheduling of the pilot and main study questionnaires. In four of the studies, the researchers piloted their questionnaires (Ely, 1986a, 1986b; Gardner & Smythe, 1981; Roger, Bull, & Fletcher, 1981). Piloting a questionnaire improves the internal consistency of the scores (Chatfield, 1988; Nunan, 1992; Seliger & Shohamy, 1989) because items that should be removed or revised are identified at relatively early stages in the studies.

Contexts. Some of the more researched target languages and contexts for SL learning have been sampled (Table 1). The largest number of questionnaires (five) come from studies done in Canada with students learning French or English in bilingual or unicultural settings (Clement & Kruidenier, 1983; Clement et al., 1976; Gardner & Smythe, 1981;

Gliksmann et al., 1982; Labrie & Clement, 1986). Three of the questionnaires were administered to students in the US learning Spanish or Japanese (Ely, 1986a, 1986b; Samimy & Tabuse, 1992). Two questionnaires were used in European settings, one in Hungary with EFL learners (Dornyei, 1990), and the other in the UK, with learners of French (Roger et al., 1981). A questionnaire given to Chinese EFL students in Hong Kong (Pennington & Yue, 1994) was a questionnaire that had been administered 14 years earlier (Pierson et al., 1980).

Constructs. The constructs being measured cover a broad range of attitudes and orientations towards a language (Table 1, see Appendix for details), but all of the selected instruments contain items intended to measure motivation, attitudes, or both. Since administering a questionnaire requires considerable planning, researchers frequently streamline the process by adapting items, measures, or entire questionnaires, from other studies. One of the selected questionnaires (Clement & Kruidenier, 1983) is a combination of measures adapted for Canadian SL learners from six previous studies (Burstall, Jamieson, Cohen, & Hargreaves, 1974; Carroll, 1975; Chihara & Oller, 1978; Gardner & Smythe, 1975; Lukmani, 1972; Spolsky, 1969). Items from Clement and Kruidenier's questionnaire were, in turn, adapted for Hungarian EFL learners by Dornyei (1990), who also adapted items from three other studies (Gardner et al., 1985; Pierson et al., 1980; Roger et al., 1981). All of the items in a questionnaire given to university students of Japanese (Samimy & Tabuse, 1992) were borrowed from a questionnaire (Ely, 1986b) which had been designed for use with university students of Spanish.

Reliability and validity. Only half of the selected studies reported estimates of reliability of scores on their motivation and attitude measures. In four of the studies (Dornyei, 1990; Ely, 1986a, 1986b; Labrie & Clement, 1986), reliability was estimated in terms of internal consistency (Chronbach alpha). In one of the studies (Clement et al., 1976), it was given in terms of Kuder-Richardson formula 20 (K-R20), and in another (Gardner & Smythe, 1981), it was reported in terms of median reliabilities).

A common error is to change a questionnaire and not re-estimate the reliability of scores obtained with the new instrument and reevaluate judgments of validity (Seliger & Shohamy, 1989). For example, a questionnaire that was developed for ESL Mexican

American migrant workers in the US (Oller, Baca, & Vigil, 1977) was adapted for use with learners of Norwegian at the University of Bergen (Svanes, 1987). The terms "Norway" and "Norwegian" replaced "the US" and "English", but otherwise, it was the same questionnaire. Both studies reported that an "instrumental motivation", typified by an interest in learning a language to improve one's chances of finding a job, was higher among some learners than an "integrative motivation", but it is difficult to judge the findings because reliabilities were not reported for the obtained scores.

Likert Scales Used in Motivation Questionnaires

Likert scales are used on many of the SL motivation questionnaires reported in the literature. A Likert scale is operationalized as "an ordinal rating scale, typically with five points, used to find the comparative strength of some attitude or opinion in a respondent, in response to a series of assertions" (Henning, 1987, p. 193). There are advantages and disadvantages to using Likert scales for rating the responses to items on questionnaires. The advantages are their relative ease of construction and the high reliabilities reported for scores obtained by this method with both large and small numbers of learners. For example, K-R20 reliability coefficients of .75 or above were reported for scores obtained from 304 respondents on seven of the 7-point Likert scale questionnaire scales in Clement et al.'s (1976) study, and Ely (1986a) reported Chronbach alpha (α) reliability coefficients of .86 for scores from 50 respondents on seven of his 6-point Likert scale items.

The main disadvantage to Likert scales is that it is difficult to say whether the breadth of the constructs under investigation are better represented by Likert scales of any particular length. Likert scales used on questionnaires vary in length, and the reasons for choosing and consequences of using a certain number of units are not always stated. Likert scales ranged in length among the selected studies from four to seven points. Pennington & Yue (1994) manipulated the Likert scale they used, by dropping the neutral response category, to counter threats to validity which they expected might arise from a neutral response set in the scores from the learners in their study. Several others of the selected studies using Likert scales dropped the neutral response category, but the authors

did not comment on their reasons for doing so (Clement & Kruidenier, 1983; Ely, 1986a, 1986b; Dornyei, 1990; Samimy & Tabuse, 1992).

Since the level of measurement will decide which descriptive statistics to use, the researcher should decide whether to assign interval or ordinal values to the data generated by Likert scale responses. If the units of the Likert scale are treated as equal intervals and the item scores are normally distributed, computing the mean and standard deviation is appropriate (Hatch & Lazaraton, 1991). If the Likert scale units are assumed to be ordinal values, that is, they are rank-ordered, median values and the semi-interquartile range are the appropriate descriptive statistics for showing central tendency and dispersion (Jaeger, 1990).

Statistical Analyses Used with Questionnaire Data

In SL motivation research, there is no uniformly applied statistical procedure for analyzing multivariate data. To a large extent, the statistical procedure to be used in the analysis will be decided by the level of measurement of the independent variables. In recent research, the more advanced statistical analyses have been increasingly applied, although their use may not be totally warranted in a particular study. For example, a condition of many of the statistical procedures for analyzing multivariate data is that there be large numbers of subjects, or respondents, so this may be a limiting factor for studies aimed at assessing motivation at the level of a single classroom or small group of learners.

Exploratory and confirmatory procedures. Two general phases of development, the preliminary (or exploratory) analysis and the definitive (or confirmatory) analysis, describe the statistical procedures used with multivariate data (Chatfield, 1988). The first phase has to do with looking at the data, reporting the descriptive statistics, and looking over the results in an exploratory way. The data may produce, in fact, unexpected results that are significant and meaningful. If the researcher is interpreting new data for the purpose of generating potentially interesting hypotheses for later study, for example, an exploratory type of analysis, such as multidimensional scaling (MDS) or factor analysis, is appropriate (Anderson, 1986; Chatfield, 1988; Everitt, 1989).

Definitive analysis, the second phase, includes model formulation, fitting, and checking, and the statistical procedures used are confirmatory (Chatfield, 1988). In SL motivation research, the use of confirmatory analyses is somewhat rare. If the researcher is interested in establishing the presence or absence of a phenomenon already described in terms of a well-formed hypothesis, a confirmatory type of analysis, such as analysis of linear structural relationships (LISREL), is done. Confirmatory studies of the socio-educational model using LISREL have been carried out with data from a modified version of the AMTB (Gardner & Tremblay, 1994a; Kraemer, 1993; Tremblay & Gardner, in press).

There is not always a clear cut distinction between the two phases of analysis, but most of the SL motivation studies conducted have been exploratory. As mentioned earlier, the motivation construct has not yet been operationalized. Although an explanation of how to do the statistical procedures is beyond the scope of this paper, an overview of three types of analysis used in SL motivation studies, that is, MDS, LISREL, and factor analysis, is presented in Table 2. The analyses will then each be briefly discussed in turn.

Table 2
Statistical Analyses Used With Multivariate Data from Questionnaires

Analysis	Level of IV	Type of Procedure	Form of Results
MDS	Ordinal	Non-parametric	A "map" of similarities and differences between items
LISREL	Continuous	Parametric	A causal model in two parts: the measurement model the structural model
FA	Continuous	Parametric	Factors that underlie a set of items

Note. IV = Independent variable; MDS = Multidimensional scaling; LISREL = Linear structural relationships; FA = Factor analysis. Continuous variables are either interval or ordinal values.

Multidimensional Scaling (MDS)

Multidimensional scaling is a statistical procedure to analyze variables that are treated as ordinal scale values. It is an efficient and relatively robust non-parametric statistical method, and therefore, few assumptions about normal distribution of the data need to be satisfied. With MDS, a simple geometric representation, or "map" results from complex information about the similarities between large numbers of items (variables) on a questionnaire (Everitt, 1989).

After administering a questionnaire and converting the learners' responses to numerical values, the researcher checks the descriptive statistics. Correlations are calculated between the item scores, and the result is a proximity matrix, in which relationships between the questionnaire items are described in terms of "similarity", if they are highly correlated, or "dissimilarity", if they are weakly correlated (Hatch & Lazaraton, 1991). A plot of the coordinates of the proximities gives a map in which data will cluster under similarities and dissimilarities. Items that are small distances from each other should be similar; items that are large distances from each other suggests they are dissimilar (Kruskal & Wish, 1978).

Theoretically, the graph created by MDS of the data can have a single dimension (a line), two dimensions (a plane), three dimensions (normal space), or more, although more than three dimensions are not considered practical (Anderson, 1986). As the number of dimensions increase, the greater the flexibility of this technique, but as researchers allow the computer to solve for greater numbers of dimensions, they must also check to make sure that the amount of disagreement between the pictorial distances and the numerical values of their data (called "stress") continues to decrease (Anderson).

The object of MDS is to reduce the data to the least number of dimensions that will coincide with a low stress value, that way, pairing simplicity with accuracy. Choosing the proper number of dimensions is not a simple matter, especially if the best representation is a model with many dimensions. Researchers have been able to accurately fix the number of dimensions in test data, using special plots, but this technique has not yet been used with questionnaire data (Hatch & Lazaraton, 1991). So, although MDS is well suited to the treatment of Likert scale values as ordinal data, the results are not as readily

interpretable as those from parametric methods, such as LISREL and factor analysis (Chatfield, 1988).

Linear Structural Relationships

LISREL is a structural equation modeling program designed to help explain the performance predicted by a set of variables. Unlike MDS, the variables in a LISREL analysis are continuous multivariate data, at the ordinal or interval level of measurement. LISREL is a parametric method, because assumptions regarding normal distributions of the data cannot be violated (see Hatch & Lazaration, 1991).

A LISREL model consists of two parts: the measurement model and the structural model. The measurement model describes the relationship between the observed variables, called indicators, and latent variables. For example, in a LISREL model, integrative orientation questionnaire items were the indicators and language attitudes and motivation were the latent variables (Tremblay & Gardner, in press). The structural model calculates regression coefficients between the latent variables, and tests the statistical significance of the inter-relationships.

To check the consistency of the socio-educational model with variables such as causal attribution and goal setting from the psychological literature, Tremblay and Gardner (in press) used LISREL to analyze data from a modified version of the AMTB. The measurement model in the analysis included as indicators 11 attitude and motivation scales from the AMTB, 14 questionnaire scales newly written by the authors, and two measures of achievement in French. Scores on the questionnaire scales had reliability coefficients (α) ranging from .26 to .92. The scales were correlated, and statistical procedures (factor analysis) showed their relationships with the latent variables, among them, language attitudes and motivation.

The LISREL structural modeling component was the basis for proposing a modified version of the socio-educational model. The LISREL model includes three moderating variables between the independent variables (e.g., the attitudes towards the French teacher and motivational intensity items), and the dependent variable, achievement in French. The moderating variables in the model are self-efficacy, goal salience, and a valence

component, which comprises two AMTB scales from the socio-educational model, "Desire to Learn French" and "Attitudes toward learning French" (Tremblay & Gardner, in press).

The aim of using LISREL in SL motivation studies, as in the above-mentioned example, is not only to confirm relationships between the independent and dependent variables, but also to confirm the direction of the paths between the variables. A LISREL statistical program may propose other models that better fit the data, so it is possible to check the direction of the relationships between variables that have been hypothesized with the direction proposed in more than one LISREL model (Tremblay & Gardner, in press).

Factor Analysis

Factor analysis is a statistical procedure related to correlation. Like LISREL, it is a parametric method appropriate for analyzing the large numbers of continuous variables in questionnaires. Unlike LISREL, it is exploratory and does not try to find causality. The product of the analysis is not graphic, as with MDS, but the factors found in the analysis are given expressive names. If the researcher plans to perform a factor analysis of the data, a general rule of thumb is that at least 51 more respondents than the number of variables will be needed for the study (Kim & Mueller, 1978). Although any rule of thumb regarding sample size is to a degree arbitrary, according to Cronbach and Snow (1977), sampling errors of correlations are large, and samples of 100 cases or more for each correlation are common when carrying out factor analyses.

The steps leading up to a factor analysis of the data are the same as those for other analyses: a questionnaire is filled in, the responses are scored, and the descriptive statistics are calculated to measure the variation across the learners' responses to the questionnaire items. In factor analysis, correlation coefficients are computed between mean or median scores of every pair of items, and analysis of the correlation matrix will set the number of variables that can be represented as a single factor. If the correlations between several items are high enough, most of the salient information from these items can be expressed by the respondents' scores on a single factor (Jaeger, 1990). Large numbers of

intercorrelations can be reduced to a few factors that account for much of the information from the data.

The first factor found in the original solution is the one that comprises the greatest possible number of all the correlated variables. The second factor finds the next largest number of variables, that are uncorrelated with the first grouping of variables. The third factor groups together as many as possible of the variables that have not already been accounted for by the first and second factors, and so on, until all of the variables have been accounted for. The eigenvalue of a factor shows how much of the variance in the set of original variables can be accounted for by the factor (Jaeger, 1990). At this stage, researches usually take factors with eigenvalues of less than 1.0 out of the analysis. The result of an original solution is a small number of factors that explain most of the variance in the original variables.

Factor rotation, a procedure that yields high correlations between a group of variables and one factor, and low correlations between those variables and all other variables, results in distinct factors that clearly represent a subgroup of variables (Jaeger, 1990). The correlation between a variable of a rotated factor and that factor is its factor loading. A high factor loading of a variable, therefore, suggests a high correlation between the variable (or item) and the factor. Researchers often try to name the factors after finding out which variables load most heavily on them.

Empirical research on motivation in SL learning often reports the findings in terms of factors. Distinctions in SL learning theory, for example, between an integrative and instrumental orientation, or between intrinsic and extrinsic motivation, is largely based on factor loadings of variables on factors from questionnaire data (see Crookes & Schmidt, 1991; Skehan, 1989). Factor analysis of data from pilot questionnaires is useful. Ely (1986b) used factor analysis to reduce more than 30 items on his pilot questionnaire to 24 items that were used on the questionnaire in the main study. Several of the selected studies used factor analytic procedures with questionnaire data in their main studies. Table 3 summarizes the reductive process of factor analysis on data from these studies.

Questionnaire Studies Using Factor Analysis

Two of the studies listed in Table 3 used virtually the same questionnaire (Pennington & Yue, 1994; Pierson et al., 1980). In the study by Pierson et al., 17 of the questionnaire items had factor loadings of .45 or larger, on 11 factors that explained 19% of the variance in the data. Factor 1 was the only factor made up of more than two items; six of the factors were single items. In the study by Pennington and Yue, factor analysis reduced the variables in the questionnaire to seven factors. Twenty-two of the attitude statements had factor loadings of .45 or above, and fewer of the factors contained single items. The amount of variance explained by the factors was not given in the study by Pennington and Yue.

There are several possible explanations for the differential results of the studies besides the time element. All of the respondents to the questionnaire in the Pierson et al. (1980) study were in Grade 10; those in the Pennington and Yue (1994) study were students in Grades 7-12 (see Table 1). As also noted in Table 1, there were approximately twice as many respondents in the study by Pierson et al. as in the study by Pennington and Yue. In addition, the Likert scales used for item responses in the Pierson et al. study were 5-point scales; those in the Pennington and Yue study were 4-point scales.

Table 3
Questionnaire Items and Number of Resulting Factors

Study	Number and Type of Items	Factors	Variance
Pierson, Fu, & Lee 1980	23 attitude statements	11	19%
Roger, Bull, & Fletcher 1981	30 attitude statements	7	76%
Gardner & Smythe 1981	118 attitude statements, plus French achievement items and aptitude measures	6	
Clement & Kruidenier, 1983	37 reasons for studying a language	9	81%
Dornyei 1990	15 language use items 44 attitude statements	4 7	71%
Pennington & Yue 1994	23 attitude statements	7	

Note. Percentages rounded off to the nearest whole number. Factor 1 explained 67% of the total variance. Factors 1, 2, and 3 explained 47% of the total variance.

Another of the studies in Table 3 was only moderately successful in reducing the number of variables in the questionnaire through factor analysis (Roger et al., 1981). The questionnaire consisted of 30 items. It was the modified version of a 51-item instrument piloted by the researchers, but it included variables from the original analysis that loaded on factors with eigenvalues of less than 1.0. Twelve of the items from the questionnaire used in the main study had factor loadings greater than .30 or smaller in magnitude than -.30 on Factors 1 and 2, which explained 76% of the total variance. Factors 3-7 all consisted of two or more items, but they accounted for relatively minor amounts of the total variance compared to Factor 1, which explained 67.1% (Roger et al.). One explanation for the larger amount of variance explained by the factors in this study than in the Pierson et al. (1980) study is the broader range of actor loadings accepted for the

variables by Roger et al. A range of $-.30$ to $.30$ is customary in factor analysis, but Pierson et al. had chosen a narrower range ($.45$ or larger), to interpret their results more conservatively.

Two of the studies in Table 3 were successful in reducing the number of variables on somewhat longer questionnaires (Clement & Kruidenier, 1983; Dornyei, 1990). Clement and Kruidenier got six factors from the original factor solution of the data from their 37-item questionnaire, which was administered in eight settings. After finding the factors for each group of learners separately, they computed a correlation matrix to compare all of the factors obtained in the eight settings. The second factor analysis yielded 12 factors, nine of which accounted for 80.5% of the total variance in the data (Table 3).

Dornyei's (1992) two-part questionnaire contained 15 language use items and 44 motivation/ attitude items. The language use items were reduced to four factors in the analysis. The amount of variance was not given, but all of the variables had factor loadings greater than $.30$ or smaller than $-.30$. The motivation/attitude items were the 18 variables that went into the second factor analysis. This analysis yielded seven factors that explained 70.6% of the variance. Factors 1, 2, and 3, comprising four variables each (31 items total), explained 47.2% of the total variance (Table 3). Factors 6 and 7 were single variable (and single item) factors, but they were included in the solution because they accounted for at least 5% of the total variance. Although Dornyei had used a minimum eigenvalue criterion of 1.0 to select the factors for further analysis in the first part of the study, he used a minimum 5% total variance criterion for the motivation/attitude variables, because it more accurately showed the gaps he observed in the factor solution.

The most reductive factoring of questionnaire data (Table 3) was accomplished by Gardner and Smythe (1981). Factor analysis of data from an early version of the AMTB, administered to Grade 7-11 learners of French, yielded six factors for each of the grades. Three of these factors, including the "Integrative Motive Factor", also showed a high level of congruence (an index of "matched" factors) across five grades of learners. More than 118 items assessing attitudes, achievement in French, and aptitude made up the 45 variables (questionnaire scales) of the AMTB used in this study (Table 3). The authors had administered a pilot questionnaire, but instead of using factor analysis to select items

for the main study, they looked for items with high measures of internal consistency and high median reliabilities (Gardner & Smythe, 1981).

The 7-point Likert scale items on the AMTB had median test-retest reliability coefficients of .67 to .86 (Gardner & Smythe, 1981), but the amount of variance explained was not given. Using 6-point Likert scales, Clement and Kruidenier (1983) and Dornyei (1990) were able to explain large amounts of the variance (Table 3) in data from questionnaires given to learners of English, French, or Spanish in a variety of contexts. The subjects in the study by Gardner and Smythe were Anglophone Canadian learners of French in Grades 7-11; similar subjects in the study by Clement and Kruidenier were all in Grade 11 (see Table 1). The large amounts of variance explained by the data in the Clement and Kruidenier study may be partly due to having respondents who were all about the same age.

By including variables hypothesized to be relevant to the motivation construct, and using factor analysis to reduce the data, it has been possible to assess and develop theories of SL motivation using questionnaires. The studies that were selected for discussion in this review are only a small sample of the research being conducted, but they serve to point out some of the different applications that have been used for questionnaires in SL motivation research. In recent reviews (Crookes & Schmidt, 1991); Oxford & Shearin, 1994; Dornyei, 1994), authors have called for more empirical evidence to augment SL motivation theories. Although it is recommended that SL motivation research and construct validation include multiple methods of assessing the motivation construct (Messick, 1989), the scope of this review is limited to using questionnaires in SL classroom settings. It is reasonable to consider how to proceed.

ISSUES FOR TEACHERS AND RESEARCHERS

Assessing the motivation of SL learners with questionnaires centers around two related issues: (a) Can existing research instruments be successfully adopted, as they are, for use in other contexts, or can they be appropriately modified for use? and (b) If existing questionnaires are unsuitable, what is the process for developing one to use for assessing motivation in SL learners?

Adopting or Adapting Questionnaires

It has been suggested that adopting an existing questionnaire to conduct further studies is the best choice. From a research perspective, replication studies play an essential role in confirming hypotheses. But finding a questionnaire relevant to the domain of interest that is also suitable in terms of the context for learning, the types of learners, and the target language will not be easy. It has been suggested, for example, that attitude questionnaires, being self-report measures, may not always show a consistency of responses when used in different cultural settings (R. W. Schmidt, personal communication, May 14, 1995). Questionnaires are developed for specific contexts and learners, and modified theories make even the most popular questionnaires go out of date. Making only superficial changes to a questionnaire that has been developed for use with other types of learners, in different contexts, or for a different language is generally discouraged (Gardner & Tremblay, 1994b).

Adapting a questionnaire is a good choice if it cannot be used as it is. A questionnaire intended for use in one specific context may be adapted for use in another, for example, by translating it into another language, by deleting or adding questions, or by rephrasing the questions. The reliability of the scores obtained on the modified questionnaire will need to be reassessed in the new context, and the results may offer valuable convergent evidence in support of construct validity (Messick, 1989). Questionnaires that have been developed for use with older or younger SL learners than those to be studied should be fairly adaptable if the target language is the same (Seliger & Shohamy, 1989).

In the studies reviewed in this paper, adapted questionnaires were used with varying degrees of success. Scores on an adapted version of the AMTB, which included 18 items that had not been on the original questionnaire (Gardner & Smythe, 1981), were reported to be reliable in a Francophone Canadian setting (Gliksman et al., 1982), as well as in the Anglophone Canadian setting for which the AMTB had been developed. The similarity of contexts may have led to the successful adaptation. Studies in which questionnaires were adapted for similar learners by changing the length of the Likert scales, or by modifying the instrument for use with learners of a different target language, reported less impressive findings (Pennington & Yue, 1994; Samimy & Tabuse, 1992).

Modifying a questionnaire for use in other contexts, therefore, should be considered cautiously. It is likely that some aspects of the motivation construct are not universal, so providing evidence of the content validity of questionnaire scores in a new context is important.

Developing Questionnaires

If existing questionnaires cannot be adopted or adapted, the only remaining choice for collecting data by questionnaire is to develop one. Constructing a questionnaire is a complex process that should begin with a review of recent developments in SL theory and issues in measurement. General advice about how to avoid the common pitfalls in questionnaire development (Nunan, 1992) is given in some SL research methods textbooks, but these reference works tend to either cover questionnaire development in general terms (Seliger & Shohamy, 1989) or concentrate on statistical analyses the researcher might use (Hatch & Lazaraton, 1991). Information about questionnaire item design is also available in social science research textbooks (e.g., Bailey, 1987), but issues specific to assessing motivation in L2 learners often has to be gleaned from published studies.

Setting up an existing questionnaire or constructing a new one presents considerable challenges. Questionnaires used in SL motivation studies reported in the SL literature are difficult to use with other learners. The prospect of having to construct an attitude/motivation questionnaire is daunting, but looking at the methodology used to develop other questionnaires, taken as a series of steps, is useful. If the process of developing a questionnaire can be clarified, research on SL motivation may be more productive.

THE PROCESS OF DEVELOPING A MOTIVATION QUESTIONNAIRE

Step One: Specify the Construct

The first step in developing a questionnaire is to specify the construct(s) to be measured. This usually means surveying the literature on SL motivation research and trying to clarify discrepancies in previous findings, or less commonly, formulating original theories about the ways in which motivation in SL learning or a related attitude can be

measured. Threats to construct validity can be avoided by controlling for construct under representation at this stage. Then the elements of the construct to be investigated are described and hypotheses are formulated. For example, if using the computer lab has not been considered in previous SL learning studies, it might be included as one of the elements of a motivation study, and using the computer lab might be hypothesized to be related to motivation to learn the language.

Step Two: Identify the Target Learners

Identifying the target group of learners is the next logical step in the process, since both the context for language learning and the ages of the respondents are important considerations in the design of the questionnaire. In the previous example, it is assumed that the learners have a computer lab that is readily available. Another consideration is the number of learners who are available to take the questionnaire, since many of the statistical procedures used to analyze the data require fairly large number of respondents.

Step Three: Accumulate Questionnaire Items

The third step is to gather items, in the form of assertions, about the construct to be studied. Sources of the items include items generated by the investigators and student and teacher interviews and open-ended surveys. It is quite common to adapt suitable parts (questionnaire scales or items) of other instruments for use in other studies. Entire questionnaires have also been adapted for use in other studies, and that way, doing away with the need to gather items, but this practice is generally not recommended. If the adapted questionnaire needs to be translated into the native language of the learners, the translations should be checked for clarity or fidelity to the assertions that are intended. Foil items, that is, items used to prevent the respondents from guessing the purpose of the questionnaire (J. D. Brown, personal communication, Oct. 13, 1995; Ely, 1986b), might also be included, to find links to undiscovered aspects of a construct, and to counter threats to construct-irrelevant "easiness" variance in the item scores. The goal of this step is to accumulate a large item pool for review and selection.

Step Four: Critique the Items

Once the items are found, the researchers (and language specialists, if possible) review the items for potential use in the new questionnaire. At this stage, redundant, ambiguous, stereotypical, or otherwise unsuitable questions are discarded. A similar process is likely to have happened when an existing questionnaire was developed, and this extra step is not always necessary when adapting questionnaires. It is still important for the researcher to verify how clear the adapted items are in their new context, since interpretation of the same questions by different learners may vary considerably with cultural background, level of education, age, and socioeconomic status, and this could contribute to construct-irrelevant "difficulty" variance in the scores.

Step Five: Choose a Rating Scale

After choosing the items, the format of the scale for the responses needs to be fixed. High reliabilities have been reported for scores in many studies using Likert scales. Thousands of respondents, therefore, must have found it relatively easy to place their attitudes or opinions along an attitude continuum, but a variety of formats might be tried on a pilot questionnaire to determine which ones are best suited to the learners of interest.

If a Likert scale is used, the sample size, characteristics of the learners, and reasonable expectations about the number of units needed to capture the range of responses all need to be considered in choosing a length. On the other hand, a large number of options on a rating scale does not necessarily make it easier to use. In addition, learners may look for (and need) a category in which to place neutral responses, despite the strategies of questionnaire developers for discouraging what might be termed fence-sitting. After the length of the Likert scale is decided on, the response categories for the corresponding points along the scale (strongly agree, slightly agree, agree, etc.) can be set.

Step Six: Assemble the Questionnaire

Assembling the items with their response scales to create a questionnaire is the next step. In the studies reviewed in this paper, many of the researchers reworded items on their questionnaires so that agreement with half of the items and disagreement with the

other half would show a consistent, favorable attitude to the assertions presented. This not only helps to deter a method effect, but also offers a means of checking for constancy in the learners' reported attitudes. Items randomly mixed for presentation on the questionnaire are favored in some studies (e.g., Ely, 1986a; Gliksman et al., 1982; Pennington & Yue, 1994); items organized as a series of scales are favored in others (Ely, 1986b; Samimy & Tabuse, 1992).

Step Seven: (Optional?) Pilot the Questionnaire

Now the questionnaire is ready for piloting. Not all researchers consider this necessary, but it reflects careful methodology. It has been suggested that many of the published SL studies would have benefited from doing a pilot study first (Crookes, 1991). The investigator gains valuable feedback and has a chance to detect flaws which may have been overlooked in a questionnaire's development. Pilot studies reviewed in this paper were conducted with as few as 50 to as many as 552 learners, suggesting that even a small scale piloting of an instrument with only 50 learners is possible, and that piloting on a large scale, with more than 500, can be done as part of the validation process.

Another advantage of piloting the new questionnaire is that the scores will help verify whether the Likert scale is an appropriate length for discriminating between the item responses. It will also show whether the scale is working with learners who are similar to the intended respondents. Computing correlations between the item scores and the total score will help identify which items have the greatest amount of discriminatory power. The distribution of the item scores on the pilot questionnaire should always be checked for normality or outliers.

One drawback to piloting the new questionnaire is that the respondents will be ineligible for the main study. This is necessary to ensure the reliability of the results obtained in the main study. Another consideration is the time element. If piloting is done at the beginning of the academic year, the analysis must proceed quickly for the improved version to be administered early in the same term. Time constraints will also decide the feasibility of setting up studies to reproduce results or of doing longitudinal studies.

The pilot study has one overriding advantage that outweighs the minor disadvantages. If some part of the experimental design is flawed, steps can be taken at a relatively early stage in the study to avoid collecting data that no amount of statistical manipulation can salvage.

Step Eight: (If the Questionnaire is Piloted) Analyze the Data

The next step is to analyze the results of the pilot study and to suggest any improvements that can be made. The questions and issues addressed in the pilot study may help clarify aspects of the first motivation study before the final questionnaire is developed. Since the results of the pilot study serve as the basis for choosing the "best" items, the final questionnaire is usually a shorter, more effective instrument than the pilot.

Steps Nine and Ten: Administer the Questionnaire and Analyze the Results

In Step Nine, the final questionnaire, which may incorporate the suggested improvements with the best items from the pilot, is administered to a large number of learners. They should be the same type of respondents as those in the pilot study, if a pilot questionnaire has been administered. In Step Ten, the results of the questionnaire are analyzed and interpreted. If factor analysis is used, specific variance can be identified and names for new or unique factors can be assigned. For the other types of analyses mentioned in this review, the results will be a map of similarities and differences with MDS, and a causal model of the factors with LISREL.

Step Eleven: (Optional) Plan Further Studies

The next step includes making plans for test-retest reliability studies and scheduling dates for administering questionnaires in longitudinal studies. At present, factor analysis and LISREL modeling can only be done on a main frame computer, so it may be necessary to arrange for time to use it. Scheduling appointments with statistical methods specialists to assist in running the programs will also be decided by the needs of the investigators.

CONCLUSION

A system of developing a questionnaire for assessing SL motivation can be discerned from the methodologies described in the twelve studies reviewed. This method is summarized as follows:

Steps for Developing a Questionnaire

- Step 1: Specify the construct and formulate hypotheses
- Step 2: Identify the target group of learners
- Step 3: Accumulate a large pool of questionnaire items
- Step 4: Critique the items
- Step 5: Choose a format for the rating scale
- Step 6: Assemble the items into a questionnaire, balancing positively and negatively scored items
- Step 7: (optional?) Pilot the questionnaire
- Step 8: (If the questionnaire is piloted) Analyze the results and interpret the findings
- Step 9: Develop and administer the final questionnaire
- Step 10: (Re)analyze the results and interpret the findings
- Step 11: (optional) Plan further studies

These steps are guidelines, rather than a prescribed series of procedures to follow, but they reasonably resemble the suggestions of experienced motivation questionnaire "consultants". Intermediate levels of the process are not included here, because they were preferences of individual researchers beyond the basic 11 steps used in all of the studies.

Steps 7, 8, and 11 are actually less optional than might be assumed. Steps 7 and 8 deal with the piloting of a questionnaire. The information gathered at these steps allows the researcher to modify and improve the questionnaire before staking everything on the results of a single administration with the target learners. Items which have been adapted for use can be looked over critically before they have a chance to adversely affect the rest of the questionnaire. The respondents should be encouraged to write their comments on the pilot questionnaire, because this will give the researcher feedback on the instrument

which might otherwise be lacking. Step 11 includes planning for confirmation of reliability and longitudinal studies. Although these types of studies have been optional in the past, renewed interest in defining the motivation construct suggests this step will be a matter of routine in future motivation studies.

REFERENCES

- Anderson, A. M. (1986). Multidimensional scaling in product development. In R. J. Brook, G. C. Arnold, T. H. Hassard, & R. P. Pringle (Eds.), *The fascination of statistics* (pp. 103-110). New York: Marcel Dekker.
- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, 80, 260-267.
- Bachman, L. F. (1990). Communicative language ability. In *Fundamental considerations in language testing*. (pp. 81-110). Oxford: Oxford University Press.
- Bailey, K. D. (1987). *Methods of social research*. (3rd ed.) New York: The Free Press.
- Bourhis, R. Y., Giles, H., & Rosenthal, D.* (1981). Notes on the construction of a "Subjective Vitality Questionnaire" for ethnolinguistic groups. *Journal of Multilingual and Multicultural Development*, 2, 145-155.
- Brown, H. D. (1990). M & Ms for language classrooms? Another look at motivation. *Georgetown University Round Table on Languages and Linguistics 1990*. (pp. 383-393). Washington, D. C.: Georgetown University Press.
- Burstall, C., Jamieson, M., Cohen, S., & Hargreaves, M.* (1974). *Primary French in the balance*. Windsor, UK: NFR Publishing.
- Carroll, J. B.* (1975). *The teaching of French as a foreign language in eight countries*. New York: Wiley.
- Chapelle, C., & Douglas, D. (1993). Foundations and directions for a new decade of language testing. In D. Douglas & C. Chapelle (Eds.), *A new decade of language testing research*. (pp. 1-22). Alexandria, VA: TESOL.
- Chatfield, C. (1988). *Problem solving: A statistician's guide*. London: Chapman and Hall.
- Chihara, T., & Oller, J. W.* (1978). Attitudes and attained proficiency in EFL: a sociolinguistic study of adult Japanese speakers. *Language Learning*, 28, 55-68.

* Source of items or scales used in questionnaires, shown in Appendix.

- Clement, R., Dornyei, Z., & Noels, K. A. (1994). Motivation, self-confidence, and group cohesion in the foreign language classroom. *Language Learning*, 44, 417-448.
- Clement, R., & Kruidenier, B.* (1983). Orientations in second language acquisition: I. The effects of ethnicity, milieu, and target language on their emergence. *Language Learning*, 33, 273-291.
- Clement, R., Smythe, P. C., & Gardner, R. C. (1976). Echelles d'attitudes et de motivations reliees a l'apprentissage de l'anglais, langue seconde. *Canadian Modern Language Review*, 33, 5-26.
- Cronbach, L. J. (1988). Five perspectives on the validity argument. In H. Wainer & H. Braun (Eds.), *Test validity*. (pp. 3-17). Hillsdale, NJ: Lawrence Erlbaum and Associates.
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington Publishers.
- Crookes, G. (1991). Power, effect size, and second language research. Another researcher comments... *TESOL Quarterly*, 25, 762-765.
- Crookes, G., & Schmidt, R. (1991). Motivation: Opening the research agenda. *Language Learning*, 41, 469-512.
- Dornyei, Z. (1990). Conceptualizing motivation in foreign-language learning. *Language Learning*, 40, 45-78.
- Dornyei, Z. (1994). Motivation and motivating in the foreign language classroom. *The Modern Language Journal*, 78, 273-284.
- Duda, J. L., & Nicholls, J. G. (1992). Dimensions of achievement motivation in schoolwork and sport. *Journal of Educational Psychology*, 84, 290-299.
- Ely, C. M.* (1986a). Language learning motivation: A descriptive and causal analysis. *The Modern Language Journal*, 70, 28-34.
- Ely, C. M.* (1986b). An analysis of discomfort, risktaking, sociability, and motivation in the L2 classroom. *Language Learning*, 36, 1-25.
- Evans, N. J., & Jarvis, P. A. (1986). The group attitude scale: A measure of attraction to group. *Small Group Behavior*, 17, 203-216.

- Everitt, B. S. (1989). Multivariate analysis: Introduction. In S. Conrad (Ed.), *Assignments in applied statistics* (pp. 108-112). Chichester, UK: John Wiley & Sons.
- Gardner, R. C. (1980). On the validity of affective variables in second language acquisition: Conceptual, contextual, and statistical considerations. *Language Learning*, 30, 255-270.
- Gardner, R. C., Clement, R., Smythe, P. C., & Smythe, C. L.* (1985). *The attitude/motivation test battery: Revised manual*. Research Bulletin No. 15. London, Ontario: University of Western Ontario, Language Research Group.
- Gardner, R. C., & Glikzman, L. (1982). On "Gardner on affect": A discussion of validity as it relates to the attitude/motivation test battery: A response from Gardner. *Language Learning*, 32, 191-200.
- Gardner, R. C., & Smythe, P. C.* (1975). *Second language acquisition: A social psychological approach*. Research Bulletin No. 149. London, Ontario: University of Western Ontario, Department of Psychology.
- Gardner, R. C., & Smythe, P. C. (1981). On the development of the attitude/motivation test battery. *The Canadian Modern Language Review*, 37, 510-525.
- Gardner, R. C., & Tremblay, P. F. (1994a) On motivation, research agendas, and theoretical frameworks. *The Modern Language Journal*, 78, 359-368.
- Gardner, R. C., & Tremblay, P. F. (1994b) On motivation: Measurement and conceptual considerations. *The Modern Language Journal*, 78, 524-527.
- Glikzman, L., Gardner, R. C., & Smythe, P. C. (1982). The role of the integrative motive on students' participation in the French classroom. *The Canadian Modern Language Review*, 38, 625-647.
- Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Boston: Heinle & Heinle.
- Henning, G. (1987). *A guide to language testing*. Cambridge, MA: Newbury House.
- Jaeger, R. M. (1990). *Statistics: A spectator sport*. (2nd ed.). Newbury Park, CA: Sage Publications.

- Kim, J.-O., & Mueller, C. W. (1978). *Factor analysis: Statistical methods and practical issues*. Beverly Hills, CA: Sage Publications.
- Kraemer, R. (1993). Social psychological factors related to the study of Arabic among Israeli high school students. A test of Gardner's socioeducational model. *Studies in Second Language Acquisition*, 15, 83-105.
- Kruskal, J. B. & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: Sage Publications.
- Labrie, N., & Clement, R. (1986). Ethnolinguistic vitality, self-confidence and second language proficiency: An investigation. *Journal of Multilingual and Multicultural Development*, 7, 269-281.
- Lazaraton, A. (1991). Power, effect size, and second language research. A researcher comments... *TESOL Quarterly*, 25, 759-762.
- Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. Malaysia: Longman.
- Lukmani, Y. M.* (1972). Motivation to learn and language proficiency. *Language Learning*, 22, 261-273.
- Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed.). (pp. 13-103). New York: Macmillan.
- Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.
- Oller, J. W., Jr. (1981a). Can affect be measured? *International Review of Applied Linguistics*, 19, 227-235.
- Oller, J. W., Jr. (1981b). Research on the measurement of affective variables: Some remaining questions. In R. W. Anderson (Ed.), *New dimensions in second language acquisition research*. Rowley, MA: Newbury House.
- Oller, J. W., Jr. (1982). Gardner on affect: A reply to Gardner. *Language Learning*, 32, 183-189.
- Oller, J., Baca, L., & Vigil, F. (1977). Attitudes and attained proficiency in ESL: A sociolinguistic study of Mexican Americans in the Southwest. *TESOL Quarterly*, 11, 173-183.

- Oller, J., Hudson, A., Liu, P. F.* (1977). Attitudes and attained proficiency in ESL: a sociolinguistic study of native speakers of Chinese in the United States. *Language Learning*, 27, 1-27.
- Oller, J. W., Jr., & Perkins, K. (1978a). Intelligence and language proficiency as sources of variance in self-reported affective variables. *Language Learning*, 28, 85-97.
- Oller, J. W., Jr., & Perkins, K. (1978b). A further comment on language proficiency as a source of variance in certain affective measures. *Language Learning*, 28, 417-423.
- Oppenheim, A. N. (1966). *Questionnaire design and attitude measurement*. New York: Basic Books.
- Oxford, R., & Shearin, J. (1994). Language learning motivation: Expanding the theoretical framework. *The Modern Language Journal*, 78, 12-28.
- Pennington, M. C., & Yue, F. (1994). English and Chinese in Hong Kong: Pre-1997 language attitudes. *World Englishes*, 13, 1-20.
- Pierson, H. D., Fu, G. S., & Lee, S.* (1980). An analysis of the relationship between language attitudes and English attainment of secondary students in Hong Kong. *Language Learning*, 30, 289-316.
- Ramage, K. (1990). Motivational factors and persistence in foreign language study. *Language Learning*, 40, 189-219.
- Roger, D., Bull, P., & Fletcher, R.* (1981). The construction and validation of a questionnaire for measuring attitudes towards learning foreign languages. *Educational Review*, 33, 223-230.
- Samimy, K. K., & Tabuse, M. (1992). Affective variables and a less commonly taught language: A study in beginning Japanese classes. *Language Learning*, 42, 377-398.
- Schumann, J. H. (1978). The acculturation model for second-language acquisition. In R. C. Gingras (Ed.), *Second-language acquisition & foreign language teaching*. Arlington, VA: Center for Applied Linguistics.
- Seliger, H. W., & Shohamy, E. (1989). *Second language research methods*. Oxford: Oxford University Press.
- Skehan, P. (1989). *Individual differences in second language learning*. London: Edward Arnold.

- Spolsky, B.* (1969). Attitudinal aspects of second language learning. *Language Learning*, 19, 271-283.
- Strong, M. (1984). Integrative motivation: Cause or result of successful second language acquisition? *Language Learning*, 34, 1-14.
- Svanes, B. (1987). Motivation and cultural distance in second-language acquisition. *Language Learning*, 37, 341-359.
- Tremblay, P. F., & Gardner, R. C. (in press). Expanding the motivation construct in language learning. *The Modern Language Review*.
- Upshur, J. A., Acton, W., Arthur, B., & Guiora, A. Z. (1978). Causation or correlation: A reply to Oller and Perkins. *Language Learning*, 28, 99-104.

Appendix

Study 1: Clément, Smythe, & Gardner, 1976

Subjects: 153 Grade 10 and 151 Grade 11 Francophone ESL students in two schools in Montreal

Constructs measured: Need Achievement, Ethnocentricity, Anxiety, Attitudes towards English Speakers, Interest in Foreign Languages, Integrative Orientation, Instrumental Orientation, Parental Influence, Anomie, Motivational Intensity, Self, Course, and Teacher Evaluations, Attitude towards Bilingualism, and Anxiety of Use in Real Situations

Description of Instrument:

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
110 attitude statements	Literature survey (Gardner & Smythe, 1975) Authors	Degree of endorsement	7-point Likert; series of scales	7 scales with K-R 20 reliability coefficients of .75 or above*
200 evaluative adjective pairs	Literature survey	Point on a continuum	Semantic differential	10 scales with K-R 20 reliability coefficients of .75 or above
20 motivation item stems	Literature survey	Multiple choice	Four options	2 scales with K-R 20 reliability coefficients of .75 or above

Likert Questionnaire Scales

- * English Class Anxiety [5 items]
- * Attitudes towards English Canadians [10 items]
- * Interest in Foreign Languages [10 items]
- * Integrative Orientation [4 items]
- * Parental Encouragement [10 items]
- * Attitudes towards Learning English [10 items]
- * Attitudes towards Americans [10 items]

Study 2: Pierson, Fu, & Lee, 1980

Subjects: 466 Grade 10 Chinese students of EFL from 8 English-medium, 3 Chinese-medium schools in Hong Kong; two subsidized schools, 3 private, 3 grant, and 3 government schools.

Constructs measured: Attitudes and achievement

Description of Instrument:

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
23 attitude statements	Author	Degree of endorsement	5-point Likert; items mixed	Factor analysis: 11 factors
50 item cloze passage	Literature survey (Oller et al., 1977)	Fill-in	Exact word	Stepwise regression with attitude factors: 6 factors (A1- A6) explained 44% of variance
20 attributes, each applied to self, an ideal, Chinese people, and Westerners	Literature survey (Spolsky, 1969)	Degree of agreement	5-point Likert	Stepwise regression with cloze scores: 5 variables for self, 4 for ideal, 4 for Chinese, 5 for Westerners

Factors

- Factor 1: Positive Orientation towards English [3 items]
- Factor 2: Desire to Converse with Westerners [1 item]
- Factor 3: Discomfort about Chinese Speakers using English (A5) [2 items]
- Factor 4: Approbation for using English (A4) [2 items]
- Factor 5: Freedom of Language Choice (A1) [2 items]
- Factor 6: English as Detracting from Cultural Identity [2 items]
- Factor 7: Self-Confidence in using English [1 item]
- Factor 8: English as a Mark of Education (A6) [1 item]
- Factor 9: Desire to Learn English (A2) [1 item]
- Factor 10: Lack of Self-Confidence in using English (A3) [1 item]
- Factor 11: Mother Tongue Favored over English [1 item]

Study 3: Roger, Bull, & Fletcher, 1981

Subjects: 86 students of French as a foreign language, 13-14 years old, in the pilot study at a school in York, UK

938 students of French as a foreign language, 13-14 years old, in the main study at several schools around Leeds and York, UK

Constructs measured: Attitudes and motivation

Pilot Study

Description of Instrument:

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
51 attitude statements	Authors Literature survey Insights of specialists	Degree of endorsement	5-point Likert; items mixed	30 items retained; 7 factors extracted. Factor 1 = 66.6% of total variance, "Attitude Towards Learning French"
1 enthusiasm scale + 1 attentiveness scale	Teachers	Point on a continuum	5-point Likert	Scores on the two scales were summed for a composite score ANOVA: students rated highly held more positive attitudes towards learning French

Main Study

Description of Instrument:

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
30 attitude statements	Pilot study	Degree of endorsement	5-point Likert; items mixed	7 factors; Factors 1 and 2 contained a few different items than in the pilot study. Factor 1 = 67.1% total variance

Factors

Factor 1: Attitude towards Learning French [8 items]

Factor 2: Utility of French [4 items]

The following had eigenvalues less than 1.0:

Factor 3: Parental Attitudes [4 items]

Factor 4: Generally Positive Attitudes towards School [2 items]

Factor 5: Attitudes towards France and French People [3 items]

Factor 6: unnamed [6 items]

Factor 7: Positive Attitude towards Language Learning [3 items]

Study 4: Gardner & Smythe, 1981

Subjects: 552 students of French as a second language in Grades 7 to 11 at one school in Canada for the pilot study;

1,521 students of French as a second language in Grades 7 to 11 from 3 schools in Canada for the second study

Constructs measured: Attitudes, motivation, aptitude and French achievement

Pilot Study**Description of Instrument:**

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
200 attitude statements	Authors Literature survey	Degree of endorsement	Likert	118 items with Median reliabilities over .50 retained
41 motivation item stems	Authors	Multiple choice	Options	24 items with Median reliabilities over .50 retained
166 evaluative adjective pairs	Authors Insights of specialists	Point on a continuum	Semantic differential	No statistical analyses: 16 items removed

Item-total correlations were computed for each grade; items in top half of at least 3 of the 5 grades were kept for the second study.

Likert Questionnaire Scales

Attitudes towards French Canadians

Attitudes towards European French People

Interest in Foreign Languages

Ratings of Integrative Orientation

Ethnocentrism

Authoritarianism

Anomie

Attitudes towards Learning French

French Class Anxiety

General Classroom Anxiety
Need Achievement
Ratings of Instrumental Orientation
Machiavellianism
Parental Encouragement

Multiple Choice Questionnaire Scales

Motivational Intensity
Desire to Learn French
Orientation Index

Second Study

Description of Instrument:

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
118 attitude statements	Pilot study	Degree of endorsement	Likert	Median reliabilities confirmed for all but one measure; 45
French achievement items (k = ?)	MLAT Students Teachers CATF	receptive, productive self rating ratings receptive, productive	norm-referenced not stated not stated not stated	variables for each grade were factor analyzed separately: 6 factors were obtained; 3 were highly congruent across the five grades. Composite
Aptitude measures (k = ?)	IQ scores	receptive, productive	norm-referenced	scale intercorrelations showed motivation composite to be a
8 evaluative adjective pairs	Authors	Point on a continuum	Semantic differential	better predictor of French achievement than the attitude or integrativeness composite.

Factors

School French Achievement
Integrative Motive
Self-Perceptions of French Proficiency

Study 5: Gliksmann, Gardner, & Smythe, 1982

Subjects: 149 students enrolled in French classes in Grades 9, 10, and 11 at a school in Ontario, Canada

Constructs measured: Attitudes, integrativeness, motivation, desire to learn French, classroom participation, achievement in French

Description of Instrument:

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
83 statements of fact and opinions	AMTB (Gardner & Smythe, 1975)	Degree of endorsement	7-point Likert; items mixed	Students were integratively or non-integratively motivated, according to median scores on 4 scales*, part of a composite Integrative Motive score
21 motivational item stems	AMTB	multiple choice	4 options	6 scales*, which contributed to composite Integrative Motive score. Justification: Lots of variance in common
classroom participation behavior	6 observations by 2 observers	Volunteering, elicited response, total correct, total incorrect	frequencies	ANOVA: Integratively motivated Ss volunteered more and gave more correct answers; no difference in elicited response freq.'s for Integ. vs. Non-Integ. groups. Patterns maintained throughout course.

Likert Questionnaire Scales

Parental Encouragement [10 items]

Need Achievement [10 items]

Degree of Instrumentality [4 items]

Interest in Foreign Languages [10 items]

Ethnocentrism [10 items]

French Class Anxiety [5 items]

* Questionnaire scales contributing to the composite Integrative Motive score are given in italics below

Likert Questionnaire Scales

- * *Attitudes toward French Canadians* [10 items]
- * *Degree of integrativeness* [4 items]
- * *Attitudes toward Learning French* [10 items]
- * *Attitudes toward European French* [10 items]

Multiple Choice Questionnaire Scales

- * *Motivational Intensity* [10 items]
 - * *Desire to Learn French* [10 items]
- Orientation Index [1 item]

Study 6: Clément & Kruidenier, 1983

Subjects: 871 Grade 11 Anglophone and Francophone students of official and minority languages in unicultural and multicultural settings in three settings, Quebec City, Ottawa, and London, Ontario, in Canada

Constructs measured: Reasons for studying a second language, Integrative Orientation, Instrumental Orientation

Description of Instrument:

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
37 reasons for studying a language	Literature survey (Burstall et al, 1974; Carroll, 1975; Chihara & Oller, 1978; Gardner & Smythe, 1975; Lukmani, 1972; Spolsky, 1969)	Degree of endorsement	6-point Likert; items mixed	6 factor solution computed for the 8 groups of students = 48 factors. Factor loadings were standardized to yield a 48 X 48 correlation matrix. Factor analysis: 12 factors; 9 factors = 80.5% of the total variance. Construct validity of integrative orientation was not supported.

Factors

- Factor 1: General Instrumental Dimension
- Factor 2: General Travel Dimension
- Factor 3: General Friendship
- Factor 4: Familiarity-Involvement
- Factor 5: General Knowledge
- Factor 6: Dominance-Recognition
- Factor 7: Anglophone Pragmatic Control
- Factor 8: Social-Cultural
- Factor 9: Curricular Language Importance

Study 7: Ely, 1986a

Subjects: 50 students in first year, third-quarter, and second year Spanish classes at a university in California in the pilot study

75 students in first year, first or second quarter, Spanish classes at that university, for the main study

Constructs measured: Language learning motivation

Pilot Study**Description of Instrument:**

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
Reasons for studying Spanish	Students in Spanish classes	open ended	none	184 items, reduced to 17 categories identified by the researcher
strength of motivation items	Students in Spanish classes	not stated	none; items	Item analysis: 7 items Chronbach alpha= .86 Factor analysis: 3 factors with eigenvalues > 1.0

Main Study**Description of Instrument:**

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
17 reasons for studying Spanish	Pilot study	Degree of agreement	4-point Likert; items mixed	3 motivational clusters; Alphas= .67, .70, .83
3 strength of motivation items	Pilot study	Degree of agreement	6-point Likert; items mixed	Regression analysis: 2* of 3 motivational clusters predicted strength of motivation

*Questionnaire scales that predicted strength of motivation shown below in italics

Likert Questionnaire Scales

Motivation Cluster A [7 items]

Motivation Cluster B [3 items]

Motivation Cluster C [2 items]

Strength of Motivation [3 items]

Study 8: Ely, 1986b

Subjects: 50 students in first year, third-quarter, and second year Spanish classes at a university in California in the pilot study

75 students in first year, first or second quarter, Spanish classes at that university, for the main study

Constructs measured: Language Class Discomfort, Risk Taking, and Sociability, and Strength of Motivation

Pilot Study**Description of Instrument:**

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
32+ attitude statements	Author	Degree of endorsement	Likert	Item analysis: 24 items, plus 4 items to assess Attitude to the Lang. Class became scale items* on the final questionnaire, plus 2 items to determine L1 use in the home and previous L2 study

Main Study**Description of Instrument:**

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
28 attitude statements	Pilot study	Degree of agreement	6-point Likert;	Chronbach alphas of 5 of 6
7 foil items	Author		series of scales	scales ranged from .65-.86.
aptitude measures	MLAT— short form	productive, receptive	norm-referenced	Author cites test developer's reliability coefficients for college students' test scores
classroom participation	4 observations by 2 observers	S seeks or gives information	frequencies	Interrater reliability: .91
oral and written proficiency in Spanish	Author Teachers	various	error counts	Oral proficiency: Pearson's $r = .98$. Written proficiency: Spearman Brown = .91-.94.

*Corresponding questionnaire scales printed in italics below

Using Questionnaires

Results (continued)

Stepwise regression: Only

Risk taking positively

predicted Class Participation.

Oral Correctness predicted by
Class Participation, Concern
for Grade, and Aptitude.

Written Correctness predicted
by Strength of Motivation and
Aptitude.

Likert Questionnaire Scales

* *Language Class Risk Taking* [6 items]

* *Language Class Sociability* [5 items]

* *Language Class Discomfort* [5 items]

* *Strength of Motivation* [7 items]

* *Attitude toward the Language Class* [4 items]

Concern for Grade [1 item]

Study 2: Labrie & Clément, 1986

Subjects: 95 Grade 9 Francophone ESL students at a Francophone high school in New-Brunswick, Canada

Constructs measured: Ethnolinguistic vitality, attitudes, motivation, self-confidence, inter-ethnic contact, communicative competence

Description of Instrument:

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
22 questions eliciting impressions comparing English and French languages and cultures	Literature survey (Bourhis, Giles, & Rosenthal, 1981)	Point on a continuum	Semantic differential	Hypothesized relationship between integrativeness and ethnolinguistic vitality not supported statistically.
16 attitude statements;	Literature survey	Degree of endorsement	Likert scale	Integrativeness minus Fear of Assimilation correlated with Relative Freq. of Contact; Motivation to Use English correlated w/ Self-Confidence, cloze test score, preference for English TV and newspapers.
20 motivation item stems;	(Clément, Smythe, & Gardner, 1976)	Multiple choice	Options	
12 language usage items;		Degree of endorsement	Options	
8 proficiency measures	Students	Self-rating	Single item	
24 frequency of contact statements	Literature survey (Prujiner et al., 1984)	Not stated	5-point additive	ANOVA showed that high quality of contact was related to greater self-confidence with English, in situations of low frequency of contact

Likert Questionnaire Scales

Attitude Towards English Canadians [8 items, $\alpha = 0.77$]

Fear of Assimilation [8 items, $\alpha = 0.52$]

Motivation to Learn English [10 items, $\alpha = 0.63$]

Motivation to Use English [10 items, $\alpha = 0.87$]

Self-Confidence with French and with English [6 items each, $\alpha = 0.76, 0.79$]

Frequency of Contact in French and in English [8 items each]

Quality of Contact in French and in English [8 items each]

Study 10: Dörnyei, 1990

Subjects: 134 Hungarian beginning and intermediate ESL students at a language school in Hungary

Constructs measured: language use intentions, beliefs, values, interests, attitudes

Description of Instrument:

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
15 language use items	Author	Relative importance to student	6-point Likert	Factor analysis: 4 factors*
44 attitude statements	Literature survey (Clément & Kruidenier, 1983; Gardner, 1985; Pierson et al., 1980; Roger et al., 1981) Author	Degree of endorsement	6-point Likert	Factor analysis: 7 factors explained 70.6% of variance, Chronbach alphas from 0.42-0.77, mean reliability = 0.60. Factors 1, 2, and 3 explained 47.2% of variance. Language Use factors highly correlated with Instrumentality, Bad Learning Experiences, and a Desire to Spend Time Abroad.

* Language Use Factors shown in italics below.

Factors

Factor 1: *Instrumental Language Use*

Factor 2: *Passive Sociocultural Language Use*

Factor 3: *Communicative Sociocultural Language Use*

Factor 4: *Reading for Nonprofessional Purposes*

Factor 1: Instrumentality

Factor 2: Need for Achievement

Factor 3: Interest in Foreign Languages and Cultures

Factor 4: Desire for Knowledge and Values Associated with English

Factor 5: Bad Learning Experiences

Factor 6: Desire to Spend Some Time Abroad

Factor 7: Language Learning is a New Challenge

Study 11: Samimy & Tabuse, 1992

Subjects: 68 University students enrolled in beginning Japanese courses in the Fall, 39 of whom were continuing students in the Spring

Constructs measured: risk taking, sociability, discomfort, motivation, attitude, grades

Description of Instrument:

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
29 attitude statements	Literature survey (Ely, 1984)	Degree of agreement	6-point Likert (?); series of scales	Stepwise regression, Fall: Female grad students taking risks are best predictors of high grades; 24.1% of variance. Stepwise regression, Spring: Strength of Motivation and Jpnse. Spoken at Home best predictors; 29.8% of variance. Motivational Clusters A and B best predictors of Strength of Motivation.
12 reasons for studying a language		Degree of endorsement	6-point Likert (?); series of scales	

Likert Questionnaire Scales (from Ely, 1984)

Language Class Risk Taking [6 items]

Language Class Sociability [5 items]

Language Class Discomfort [5 items]

Strength of Motivation [7 items]

Attitude toward the Language Class [4 items]

Concern for Grade [2 items]

Motivation Cluster A [7 items]

Motivation Cluster B [3 items]

Motivation Cluster C [2 items]

Study 12: Pennington & Yue, 1994

Subjects: 285 students of EFL in eight Hong Kong secondary schools, grades 7-12, ages 12-20.

Constructs measured: attitudes and achievement

Description of Instrument:

<u>Number and Type of Items</u>	<u>Source</u>	<u>Response Format</u>	<u>Type of Scale</u>	<u>Results of Statistical Analysis</u>
23 attitude statements	Literature survey (Pierson, et al., 1980)	Degree of agreement	4-point Likert; items mixed	4 items previously receiving highest endorsement receive very little in present study; Factor Analysis: 7 factors, disagreement with 3 of the 4 items loaded on Factor 1. Percentages of variance due to each factor not given.

Factors

Factor 1: English as Not Detracting from Cultural Identity [3 items]

Factor 2: Positive Orientation towards English [5 items]

Factor 3: Social and Instrumental Value of English [4 items]

Factor 4: Positive Orientation to Educational and Official Status of English [3 items]

Factor 5: Lack of Discomfort about Chinese Speakers using English [2 items]

Factor 6: Positive Cognitive-Affective Orientation towards English [2 items]

Factor 7: Disjunction of Views on Intrinsic Quality of English Language and English-Based Culture [2 items]

Patricia O'Bryen

Hawai'i English Language Program

1395 Lower Campus Road

Makai Campus 13-1

Honolulu, HI 96822

e-mail obryen@hawaii.edu