

3 The coding of reported emotional experiences: antecedents and reactions

Heiner Ellgring and Eva Bänninger-Huber

3.1 Introduction

One of the general aims of this study was to determine whether there are specific kinds of situations which elicit emotional reactions and what sort of non-verbal and bodily reactions are experienced in relation to these emotions. As described above (Section 2.3), open-ended questions were used to obtain as much information as possible on these issues and to avoid biasing subjects toward stereotypical or socially desirable answers. While this procedure yielded rich information, coding of the qualitative descriptions of antecedents and consequences was required to allow quantitative data analysis. In this chapter, the development, nature, and application of coding systems used for this purpose are described. This includes description of the general rationale behind the coding, of the construction principles used, and of the reliability of coding.

Investigation of antecedent situations and behavioural reactions to emotions touches two quite different areas within social psychology which have both received increasing attention in recent discussions: the concept of situation as a determinant of behaviour and the concept of non-verbal behaviour as an expression of emotion. Since both aspects each require a different approach, the coding of antecedent situations and of reactions to emotions will be described separately. In the following sections, the coding of situations as antecedents of emotions is described first, followed by the coding of non-verbal and physiological reactions. Since they were developed together, most of the construction principles apply to both codes. It should be noted that these codes cover only the free descriptions given by the subjects. Rating scales and other categories used in the questionnaire are described in Chapter 2.

3.2 Antecedent situations

Given the impact that is ascribed to social situations in various theories of personality, behaviour, and emotions, surprisingly little effort has been made to systematise it. So, when in 1979 we began the investigation of

which kinds of situations elicit emotions, we had to develop a procedure that allowed their description and quantitative comparison.

A taxonomy of emotion-eliciting situations can only partly rely on objective characteristics. Situations are evaluated and defined by the individual (McHugh, 1968). Moreover, as is discussed in Chapter 1, the subjective definition of the situation is the chief determinant of the resulting emotion. Furthermore, a classification using only a few dimensions is likely to be insufficient, since such situations are highly complex in nature.

The literature was of only partial use in relation to achieving our goal. Situational aspects are dealt with, especially in connection with 'interactional' personality theories (e.g. Endler & Magnusson, 1976; see also contributions in Magnusson, 1981). According to these theories the situation has more influence in determining cognitions and behaviours than do the traits of an individual. The social episode theory (e.g. Bowers, 1973; Forgas, 1979) sees situations as cognitive representations of events that people experience in their everyday lives. Diaries of daily events are used here, as sources of information. The problem for quantification lies in defining the unit of analysis, that is in deciding when an episode is finished or when it starts, etc. No general classification of situations has emerged from this approach up to now.

Since the completion of our analysis, some procedures that take into account different classes of emotional situations have been published. In most cases, situations are presented to the subject in order that subjective and physiological reactions may be recorded (Schwartz & Weinberger, 1980; Averill, 1982; Boucher, 1983). Quite frequently, stress or anxiety-provoking situations are used in this way (Sells, 1970; Magnusson & Stattin, 1981; McGrath, 1982). In a recent publication, Brandstätter (1983) presents data on situations which were taken from diaries of subjects. Besides being often based on small samples of respondents or restricted to stress/anxiety concepts, these approaches are mostly not representative because of content restrictions. It is not possible to tell which situational aspects are recalled as eliciting agents for the various emotions.

For emotion-related situational features, then, there have been, up to now, no generally accepted classifications available which are broad enough to cover a variety of emotions and are sufficiently empirically validated. This might be due to the lack of a general theory of situations or their relevant dimensions. A theory-guided selection or construction of situations or situational components is therefore only partly possible. At this stage in our knowledge, an inductive empirical procedure had to be applied to the situations described by the subjects in their free responses in order to classify and quantify this information.

A somewhat eclectic approach seems justified at the point where we first

have to gather information about what kinds of events are reported as eliciting discrete emotions. Some of the situational aspects clearly need to be defined on *a priori* grounds. One of these is the social aspect of the emotional experience. From a situational point of view it is important to know about other people who are involved, for instance how many are there and what is their relationship to the subject. The main question, however, concerns the content and the qualitative characteristics of the situations.

The construction of the codes

The final version of the coding scheme emerged from an interactive process whereby categories were developed on the basis of theoretical considerations and modified during the first attempts to apply the coding to the material available. The codes changed again after an extensive pilot study. Another stage of this process was the combining of categories in order to summarise the actual content. So, inductive as well as more theoretical strategies were used in the construction. The inductive procedure proved to be particularly necessary for coding the content of situational descriptions.

There were four stages of construction. First, a German sample of 20 subjects took part in a semi-structured interview. A rough content analysis of situation descriptions as a first approach to category construction was then carried out. Second, in a pilot study, 20 subjects completed a preliminary version of a questionnaire and a coding schema for situational aspects and behavioural reactions was developed. It proved necessary to add to the categories derived from the content of those categories which had not appeared yet, but which were of theoretical importance.

In a first study, described by Scherer, Summerfield & Wallbott (1983), this code was applied to 626 questionnaires of subjects from five countries (France, Great Britain, Italy, Switzerland, and West Germany). Some categories were added in order to include some country-specific aspects. For example, 'fear of supernatural events' occurred in the British sample and thus made an additional category necessary. The experience from this pilot study led to the present version of the code.

In a third stage, for the final version, the new form of the coding schema was applied to a small sample of 20 interviews. Again, the schema was expanded with a few categories which appeared necessary to cover as many elements from the answers as possible. Categories were made as comparable as possible for the different emotions. Code 01, for instance, always indicated news in the immediate social context, code 02 indicated news in the mass media, code 12 indicated situations in which strangers were involved, and so on. This kind of code applied to each of the four emotions.

The final version was then used to code the whole sample of questionnaires after reliability checks had been successfully completed. For this final

version, translations into all of the six languages that were involved in the study were available, that is French, English, Hebrew, Italian, Spanish, and German.

After the coding of all the questionnaires was completed some of the codes were combined for data analysis. A single code was put together with another one if it had a very low frequency of occurrence and was similar in content.

The situation code

For each emotion, there was a situation or 'antecedent' code containing a varying number of categories. For joy, there were 20, for sadness, 19, for fear, 25, and for anger, 21 categories. In order to allow comparisons between emotions, a set of similar categories was used for all of the four emotions. The categories 'good news' (01, 02) for joy, for example, are comparable with the categories 'bad news' (01, 02) for sadness, fear, and anger. In addition, emotion-specific codes like 'fear of traffic' (17) were used, resulting in different numbers of categories for each emotion. For each emotion, there was also a category for 'uncodable situations'. A detailed description of all coding categories used is provided in Appendix C. Coders were instructed always to code the most specific category with the least degree of inference. Double coding was allowed in complex situations where at least two elements were present in the situation itself.

The frequencies of single categories, especially for joy, fear, and anger, show markedly skewed distributions (for details see Chapter 5). For four of the respective categories, relative frequencies were 10% or more. These four categories cover more than half of the codings: 61% for joy, 57% for fear, and 68% for anger. The remaining categories were used for only 1-4% of the situation descriptions. A few of the categories rarely or never appear: 2 of the categories for joy, 3 for fear, and 5 for anger.

In the case of sadness, the situation is somewhat different. Only two of the categories had a relative frequency of 10% or more. The distribution of relative frequencies over the categories is not as skewed as for the other emotions, that is 67% of the codings are distributed over 18 categories. Thus, nearly all of the categories were used quite frequently for coding antecedent situations for sadness. It should be noted that only 1-3% of the situations were uncodable within this framework.

From these findings, the question arises as to whether the different frequency distributions result from the differing resolution power of the categories. When, for example, 'relationships with friends' occurs in 22% of cases as an antecedent of the emotion 'joy', the category might not be sufficiently specific. Rather, different situations might be being included in this aspect. This would lead to a more frequent coding of broad categories

than of specific ones. For example, the category 'relationships with friends' (especially for joy) covers a relatively broad spectrum of different emotion-releasing situations and thus results in a frequent coding.

This argument does not hold for the other emotion codes, however. For fear, for example, the category 'fear of traffic' (17) was the most frequent although it is quite a specific one. From this it might be concluded that the skewed distributions observed here are due to emotion-specific effects rather than to the insufficient resolution power of single categories.

It is not clear why a few of the categories appeared very rarely or not at all. Emotion-specific effects may play a role in the sense that these situations were of little or no relevance for these emotions.

Reliability of the final code

The coding was performed by trained coders in the respective universities. Training included practice trials with the coding procedure and discussion with the staff in order to establish inter-coder reliability. The number of paid student coders used in the various countries varied from two to six.

As an important criterion for the quality of the coding procedure, inter-coder reliability was assessed at different stages independently for situational and reaction codes.

In an 'international reliability study on antecedents', 10 randomly selected situations for each emotion, that is 40 situations, were coded by two coders in each country. Double coding was allowed. Reliability (% agreement) between the two coders in each country was established as well as agreement of the coders with a criterion coding. The latter agreement was averaged over the two coders in each country. The criterion coding was established during a meeting of the study group where 40 situation descriptions were coded and discussed. In cases of criterion double coding, agreement on both codes was scored '1', agreement on only one code was scored '1/2'. The percentage agreement between coders was calculated as:

$$p = N_{\text{agreement}} / \text{Total number of situations presented.}$$

Table 3.1 gives the corresponding figures.

Generally, coder agreement was reasonably high for the antecedent codes. The average agreement *between coders* varied very little between emotions ($p = 76\text{--}78\%$) whereas between countries the variation was somewhat stronger ($p = 63\text{--}90\%$) (Table 3. 1a). As an index of the validity of the codings with reference to a standard, the agreement of the coders with the criterion coding was determined. Values were similarly high for the average like agreement between coders with similar variability (between emotions, $p = 70\text{--}81\%$; between countries, $p = 66\text{--}86\%$). An average agreement of 76% can be considered an acceptable level of reliability for the coding of

Table 3.1. *Reliability of the antecedent code (percentage agreement between coders)*(a) *Agreement between coders in each country*

Country	Emotion				Average
	Joy	Sadness	Fear	Anger	
Belgium/France	70	70	80	60	70
Great Britain	65	80	85	70	75
Israel	90	80	75	80	81
Italy	80	70	50	50	63
Spain	70	80	90	90	83
Switzerland	75	100	85	100	90
West Germany	80	65	65	95	76
Average	76	78	76	78	77

(b) *Agreement of coders with criterion*

Country	Emotion				Average
	Joy	Sadness	Fear	Anger	
Belgium/France	80	75	60	70	71
Great Britain	70	75	65	55	66
Israel	80	68	63	73	71
Italy	80	75	70	75	75
Spain	80	70	75	75	75
Switzerland	88	85	80	80	83
West Germany	88	83	75	98	86
Average	81	76	70	75	75

situational aspects, particularly given the variability and complexity of the situation reports.

The final version of the code reflects a process during which several modifications were made. This was mainly due to the rich and complex information given by the subjects. Our goal of enabling quantitative comparisons of content to be made could have led to the definition of rather abstract classes. However, a high amount of specific content information was retained in the coding schema of the final version.

3.3 Behavioural and physiological reactions

The emotional response, besides the subjective feeling state, is accompanied by non-verbal behaviours, physiological symptoms, and/or verbal behaviour. Quite apart from theoretical views about the order of occurrence of these

reactions, they are most probably regarded by the individual as being a consequence of emotion which he or she may also attempt to control or reduce.

Non-verbal behaviour has attracted researchers as a possible indicator of discrete emotions for a long time. The expression of emotion on the face has been studied since Darwin (1872) (for a recent study see Ekman, Friesen & Ancoli, 1980). The voice has been examined by Scherer (1982) and other bodily movements have been studied by Wallbott (1982a).

Non-verbal behaviour can be measured in a variety of ways (see Scherer & Ekman, 1982). From the various methods reported in the literature it would be possible to adopt categories in order to classify behavioural reactions described by the subjects. Reports on physiological reactions may be classified in a similar way. Such a method was proposed independently after the completion of this study by Pennebaker (1982).

When describing non-verbal and physiological reactions, it should be kept in mind that the reports on these behaviours and sensations reflect only the subjectively remembered aspects which come into the consciousness of the individual. Thus, fewer behavioural changes may be reported than actually happened. It therefore seems sufficient to use relatively broad categories for these reactions rather than specific or differentiated behavioural descriptions as would emerge from applying codes similar to elements of the Facial Action Coding System (Ekman & Friesen, 1978).

The construction of the code

The procedure for coding non-verbal behaviour, physiological symptoms, and verbal behaviour, as in the case of the coding of antecedent events, has to be adapted to the material reported by the subjects. The method of categorising should be as open as possible in order to make use of all the types of information given. Similar strategies to those used for coding antecedent events were therefore used in constructing a coding schema for non-verbal behaviour and physiological reactions to emotions.

Reactions to emotional experiences were coded with regard to four general areas: (a) vocal behaviour, (b) non-verbal visible behaviour, (c) behavioural tendencies, and (d) sensations and vegetative symptoms. Within this 'symptom code', these aspects were assessed by means of categories which either could be generally applied, like 'experienced emotional quality', or were specific to the particular behaviour, like 'harsh voice'. Table 3.2 gives an overview of the code, which is described in detail in Appendix C.

As can be seen from Table 3.2 most of the categories were concerned with the description of vegetative sensations and symptoms (9), movements and postures of bodily parts (5), and behavioural tendencies (7). The area of cognitive symptoms will not be analysed further, since these symptoms were hardly ever reported.

Table 3.2. *Number of categories in the non-verbal behaviour and bodily reaction code*

Categories	
Vocal behaviour	
1. Speech	3
2. Voice	7
Non-verbal visible behaviour	
3. Facial expression	13
4. Gaze	5
5. Movement and posture of bodily parts	24
6. Body movement, displacement, and posture	10
Behavioural tendencies and intentions	
7. Behavioural tendencies	22
Sensations and vegetative symptoms	
8. General symptoms	4
9. Vegetative sensations and symptoms	34
10. Cognitive symptoms	3

For reaction classes 1 (speech) to 6 (body movement, displacement, and posture), an emotional qualifier (Experienced Subjective Quality, ESQ) could be coded, if it was mentioned explicitly. The ESQ was coded when a general description was given, such as 'my face became calm and quiet' (code 315 = decreased facial activity) or 'no muscle in my face moved' (code 316 = controlled). Coders were told, however, to check for the possibility of more specific codes (like 'smile', or 'painful smile') first. Thus, ESQ was coded for non-specific qualitative changes of specific bodily areas, for speech, and for voice. It was coded according to the following categories: normal, aroused positive or negative ('tight, nervous, tense'), increase ('fast, much, strong') or decrease ('slow, little, weak'), controlled, and changed (unspecified).

Associated emotions (happiness, sadness, etc.) could also be coded if explicitly mentioned by the subject. This was the case if, for example, a person reported that he or she reacted with a 'happy', 'angry', 'sad', etc. face, gaze, or other behaviours. For the vegetative sensations and symptoms, metaphors like 'butterflies in the stomach' were also accounted for (code 973). Multiple coding was permitted within this code.

For data analyses, a combined code was defined in order to compare the reaction tendencies across the various emotions (see Appendix C). The combination was done firstly according to low frequency and secondly according to comparable content. With these combinations the mere occurrence or mention of speech symptoms, face reactions, muscular reactions, etc. was counted without differentiating between specific codes.

The results are discussed in detail in Chapters 6 and 7. We report here only briefly on the general frequency of the different categories.

All the non-verbal reactions were distributed on 27% vocal reactions, 38% face and gaze reactions, and 27% bodily movements. Most of the vocal reactions occurred for voice (21%), especially in the case of anger. Laughing and smiling were found, as would be expected, together with joy (41%), whereas crying was mainly connected with sadness. Non-specific face reactions occurred in 15% of the situations, predominantly in association with negative emotions.

Physiological symptoms were mainly reported as muscular reactions (21%), stomach reactions (12%), a rise in blood pressure (10%), and temperature changes (10%). Non-specific unpleasant sensations were also quite frequent (14%).

Non-verbal reactions were reported on average in 65% of the situations and physiological symptoms in 81% of the situations. This may indicate some higher sensitivity with respect to physiological reactions. It is not possible to say, on the basis of the self-reported reactions, whether non-verbal reactions – as a consequence of stronger control tendencies – occur less often than physiological reactions in emotional situations.

Reliability of the symptom/reaction code

Inter-coder agreement was calculated for the detailed as well as for the combined code (see below). In order to assess reliability, 20 descriptions, 5 for each emotion, were coded independently by two coders in four countries. (Owing to organisational constraints, this reliability study could not be conducted in all countries.) The situations were chosen according to representativeness of symptoms. Reliability was assessed for each country independently.

It was to be expected that combined codes would give a higher agreement since the combination was calculated after the original coding had been done. Table 3.3 gives the reliability (percentage agreement) of the bodily reaction code for the different emotions and the different countries. Agreement was calculated according to $p = \frac{2 \times N}{I + N}$ agreements (N of coder 1 + N of coder 2) where N is the number of coded symptoms. The average agreement of the coding of non-verbal behaviour and bodily reactions ranged from 69% to 86% for the detailed code. Reactions to sadness and anger were coded with somewhat less agreement than reactions to joy and fear. Considering the complexity of the phenomena coded, inter-coder agreement appears to be sufficiently high. The variability of the average agreement between countries was low enough to warrant comparable use of the coding schema across the cultures.

Table 3.3. *Reliability of the non-verbal behaviour and bodily reaction code (percentage agreement between coders)*

Country	Emotion				Average
	Joy	Sadness	Fear	Anger	
Italy	69	81	78	68	74
Spain	91	58	100	44	71
Switzerland	86	68	93	100	86
West Germany	64	64	74	75	69
Average	78	68	86	72	75

From the general results reported in this section it can be seen that the code allows a reliable and differentiated categorising of a broad range of symptoms. The description of non-verbal and physiological reactions by the subject obviously does not tell us anything about the actual behaviour. This has to be assessed separately by observational methods and physiological measurement. The validity of the information gained in this way can, therefore, only refer to the subjective experience associated with an emotion. This, on the other hand, is important when there are attempts to control the emotion and its features.

3.4 Conclusions

The task of classifying and quantifying situational descriptions and non-verbal and physiological reactions was a difficult one given the complexity and richness of the qualitative information. Nevertheless, this task appears to have been successfully accomplished by means of the two coding systems developed here.

In order to evaluate the quality of coding systems like these, formal and content criteria have to be fulfilled. As formal criteria, the reliability and validity of the coding have to be established. In our study, the same procedures have been applied to material of a highly complex content that was gathered under quite different cultural conditions. Reliability expressed as inter-coder agreement can be regarded as being sufficiently high given the complexity of the material to be coded and the inter-cultural setting.

The validity of the coding system is much more difficult to determine. As mentioned before, subjective reports cannot be equalled to behavioural or physiological measures. However, validity can be inferred from comparisons described in the following chapters. As one criterion, values derived from these procedures should be able to differentiate between emotions and

individuals (see especially Chapters 5, 6, 7, and 10). As is described in the following chapters, the system proved to be valid in this sense.

Content criteria are even more difficult to evaluate. Our general strategy was to build up the coding system primarily in an inductive way, that is we were guided by the material given by the subjects. Experience from two preliminary studies, especially the pilot study with over 600 subjects, was used for the compilation of the final form. Thus, the broadness of contents seems to be sufficiently accounted for in the code. As a quantitative criterion, the low percentage of 'uncodable' information may also be regarded as supporting evidence. For some aspects, however, a further split of categories might yield more differentiated information. A situation code like 'meeting friends' as an elicitor of joy is probably too broad a category containing too many divergent aspects. A further task would be to define some of the categories more finely, with more examples for coding. Inter-coder reliability could be improved by refined definitions as well as by more extensive training of the coders.

Generally, the coding schemata are tools which can also be used in other studies. Translation into six languages and their examination in eight countries make them applicable in different cultural contexts. In our study, the coding system gives access to antecedents and consequences of emotions. Differing from traditional questionnaire approaches, the categories developed take up and quantify not only predetermined answer possibilities but also highly complex information given by the individual in a free response.