The Adult Attachment Interview: coherence & validation in adolescents

Beijersbergen, M.D.

Chapter 2

The Concept of Coherence in Attachment Interviews: Comparing Attachment Experts, Linguists, and Non-Experts

Beijersbergen, M.D., Bakermans-Kranenburg, M.J., & Van IJzendoorn, M.H. (2006). *Attachment & Human Development, 8,* 353-369.

Abstract

Coherence is a central construct in attachment interviews. Nevertheless, the concept has never been the main focus of a study in the attachment field. The present study examined whether coherence in attachment interviews is defined differently by experts trained in attachment theory, by linguists, and by non-experts. The 72-item Coherence Q-sort (CQS) was used to determine the profile of a prototypical coherent interview. Results indicated that attachment experts could be reliably distinguished from the (combined) other groups: attachment experts emphasized quality and manner more than all other groups, linguists emphasized quantity and relevance more than attachment experts, and higher educated non-experts valued relevance more than attachment experts. Defining coherence in attachment interviews is thus more than just applying Grice's linguistic maxims; expertise in attachment theory is critical for defining interview coherence. Consequences for the coding of the AAI by non-attachment experts, as well as computer coding (im)possibilities are discussed.

Introduction

Coherence is a central construct in attachment interviews (Main, Goldwyn, & Hesse, 2003; Waters, Treboux, Fyffe, & Crowell, 2001; Zeanah, Benoit, & Barton, 1994). An essential feature of these interviews is that participants are asked for general evaluations of relationships and/or events as well as actual evidence supporting these evaluations. Examples of frequently used interviews in the field of attachment are the Adult Attachment Interview, (AAI; George, Kaplan, & Main, 1996; Main et al., 2003), the Current Relationship Interview (CRI; Crowell & Owens, 1996), and the Working Model of the Child Interview (WMCI; Zeanah, et al., 1994). Participants can be classified as having a secure or insecure attachment representation in the AAI and CRI, and having a secure or insecure representation of their infants in the WMCI. During these interviews participants are faced with two tasks: (1) producing and reflecting upon memories related to attachment while simultaneously (2) maintaining coherent discourse with the interviewer (Hesse, 1996). Adults with a secure attachment representation are able to fluidly shift their attention between these two tasks. Hesse (1999) suggested that this flexibility of attention may be a necessary prerequisite to sensitive and responsive caregiving.

When can discourse be called coherent? The linguistic philosopher Grice (1975) formulated a general principle for rational, coherent discourse, called the Cooperative Principle: “Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged” (p. 47). Four maxims fall under this principle, namely:

Quality:	be truthful, and have evidence for what you say
Quantity:	be succinct, yet complete
Relation/Relevance:[1]	be relevant
Manner:	be clear, brief and orderly

In coherent discourse, participants adhere to these four maxims. Grice (1975) proposed that the maxims are not arbitrary conventions, but rational principles for cooperative exchanges. Conversational participants seem to assume that, ceteris paribus and in the absence of indications to the contrary, the Cooperative Principle and the maxims will be observed. However, participants may sometimes fail to fulfill a maxim: they may quietly violate a maxim, opt out, face a clash of maxims, or flout a maxim. Mura (1983) noted that violations of the maxims are legitimate when they are

[1] Grice referred to this maxim as the maxim of relation. It is however better known as the maxim of relevance. In this paper we will therefore refer to relation as the maxim of relevance.

licensed by directly appealing to Grice's Cooperative Principle or by appealing to the maxim of quality when violating one of the other three maxims. An example of licensing a violation of the maxim of quantity is "I am sorry but I would rather not go into that". Of the four maxims, Grice (1975) suggested that quality might be seen as the most important maxim. The other maxims are supposed to be applicable only on the assumption that the maxim of quality is satisfied. Grice (1975, p. 46) also noted that a part of the maxim of quantity, be succinct, is perhaps not necessary because it will be covered by the maxim of relevance.

In attachment interviews, individuals are classified on the basis of the properties of their discourse. These properties are consonant with Grice's Cooperative Principle and the four maxims (Hesse, 1999). Discourse is called coherent when the participant is able to access and evaluate memories while simultaneously remaining truthful (quality) and collaborative (quantity, relevance, and manner) (Hesse, 1996). In the AAI, secure participants only marginally violate Grice's maxims. When a speaker commits transgressions of Grice's maxims, the interview discourse is considered less coherent. It should be noted that the protocol of the AAI is suggested to have the potential of surprising the unconsciousness. Because of the relative rapid pace of the interview and the many complex questions, ample opportunities are provided to violate Grice's maxims such as by contradictions (George et al., 1996). Insecure dismissing adults typically violate the maxims of quality and quantity. These adults are not able to give evidence for the positive evaluations they provide or even contradict themselves, and they may claim lack of memory. Insecure preoccupied adults tend to make transgressions of quantity, relevance and manner. They tell long stories, drift away from the main topic of the question and use angry or passive speech. The two different forms of insecure attachment representations are thus characterized by different forms of incoherent discourse (Main et al., 2003). The importance of the coherence scale in the AAI was shown empirically by Waters and colleagues (2001) who found that the coherence of transcript scale is the most important component of an empirically derived continuous security score.

Grice's maxims, which have been applied to the study of attachment, are rooted in the field of linguistics. In linguistics, Grice's maxims have been discussed extensively almost from the beginning (see Haberland & Mey, 2002, for a review). One major question is whether it is necessary to have four maxims. Horn (1989), for example, only focuses on two principles: the Q-principle (quantity) and the R-principle (relevance). Moreover, Sperber and Wilson (1995) posit in their theory of relevance that only one maxim is needed. They argue that everything said would be guided by the principle of relevance: what people say is relevant or else they would not say it. In their view of communication people try to minimize efforts and maximize rewards when processing information. This theory has received much support (Blakemore,

1987; Carston, 1987; Kempson, 1987) as well as much criticism (Giora, 1997; Levinson, 1989; Seuren, 1988). A frequent objection against the theory of relevance is that because of its emphasis on economically rational behavior, important factors that do play a role in human communication have been excluded (Hinkelman, 1987; Mey & Tabot, 1988). Wilks (1987, p. 735) for example argued that relevance "is always to someone" and cannot be defined objectively.

Despite the discussion surrounding Grice's maxims, some or all are still used in work on linguistics. Saygin and Cicekli (2002), for example, investigated the relation between Grice's maxims and the success of computers in imitating human language use by applying a variant of the Turing Test (Turing, 1950; for a review see Saygin, Cicekli & Alkman, 2000).[2] A computer, a human being, and an interrogator are involved in this test. The interrogator stays in a separate room and needs to find out which one of the two entities he or she is conversing with is the human. Saygin and Cicekli (2002) used conversation excerpts of the interrogator and the computer. Subjects were asked whether the computer was successful in imitating human language use and whether the conversational maxims were violated. It was found that violations of the maxims of quantity and relevance revealed the identity of the computer, while manner violations were seen as human-like.

Although coherence is a central construct in attachment interviews, most studies only report on attachment classifications. For example, it has been shown that attachment representation is associated with infant's attachment security, parent's sensitivity (Hesse, 1999; Van IJzendoorn, 1995), social adjustment (Crowell et al., 1996) and psychopathology (Patrick, Hobson, Castle, Howard, & Maughan, 1994; Rosenstein & Horowitz, 1996; see Dozier, Stovall, & Albus, 1999, for an overview). Some studies also use coherency scores. Fonagy, Steele, and Steele (1991) showed that in the AAI mothers of securely attached infants had the highest coherence scores, significantly distinguishing them from mothers of avoidant infants. Dickstein and colleagues found an association between parent's coherence during family narratives and (observed as well as self-reported) family functioning (Dickstein, St. Andre, Sameroff, Seifer, & Schiller, 1999). In a study on preschool coherence, a relation was found with infants' attachment security (Sher-Censor & Oppenheim, 2004). These studies demonstrate the existence of a link between coherence and other important attachment constructs.

Some researchers compared groups of clinical and non-clinical subjects to examine, among other things, whether they showed different levels of coherence.

[2] In the original Turing Test gender was an important issue. A man and a computer had to convince the interrogator that they were women. Later work with the Turing Test mostly ignored the gender issue (Saygin, Cicekli, & Akman, 2000).

Fonagy and colleagues (1996), for example, showed that psychiatric (non-psychotic) inpatients were less able to maintain coherent discourse than case matched control subjects. Upper middle-class subjects who had been psychiatrically hospitalized at age 14, were found to be less coherent in the AAI than control subjects when they were 25 years of age (Allen, Hauser, & Borman-Spurell, 1996). Recently, Barone (2003) found that a clinical group with borderline personality disorder had dramatically lower scores on coherence than the non-clinical group. Clinical and non-clinical groups, therefore, seem to differ in coherence of discourse in the AAI.

The AAI is a labour-intensive instrument: not only because of the coding process itself, but also because of the training necessary to become a reliable coder. Computer-based linguistic content analyses might make the coding of AAIs less time consuming and more accessible for non-attachment experts. Buchheim and Mergenthaler (2000) analyzed interview transcripts of 10 dismissing, 10 preoccupied and 20 autonomous adults with a text analysis computer program. They assessed (1) *emotional tone* by measuring the proportion of word forms which express affect, (2) *abstractness* by measuring the proportion of abstract word forms, e.g., words ending in -ness, -ment, or –tion, (3) *emotion-abstraction patterns* by looking at the interaction of emotional tone and abstraction, and (4) *referential activity* on the basis of the proportions of words standing for the concreteness, specifity, clarity and imagery of a text. Significant differences among the three attachment categories were found for emotional tone and referential activity. Dismissing adults scored the lowest on both aspects, while preoccupied adults scored the highest and autonomous adults scored in between. A discriminant analysis with emotional tone and referential activity as predictors correctly predicted membership of the three attachment groups in 60 percent of the cases (Buchheim & Mergenthaler, 2000).

Appelman (2000) also applied computer-based text analyses to AAIs, assessing emotional tone, abstractness and referential activity of the AAI fragments where the subjects are asked for adjectives describing the relationship with each of their parents and for evidence supporting those adjectives. Secure respondents scored higher on emotional tone and referential activity than insecure respondents, but no differences were found for abstraction. The dismissing and preoccupied transcripts did not significantly differ from each other on any of the linguistic measures.

Computer programs thus identified differences among the attachment groups. However, the programs did not assess coherence. Buchheim and Mergenthaler (2000, p. 403-404) noted that “neither the Emotion-Abstraction Patterns nor the CRA [CRA = Computerized Referential Activity] can measure this aspect.… mere consideration of the score on the language measures within the attachment groups is not suitable as a direct substitution of a complex discourse analysis of the AAI.” They proposed that future research should focus on identifying linguistic markers for

coherence. So far no new results with respect to measuring coherence with computer programs have been reported.

Despite the importance of coherence in attachment interviews, the concept has never been the main focus of a study on attachment. The purpose of the present paper is to examine whether people of diverse backgrounds define coherence differently. The question is whether coherence as referred to by attachment experts pertains to the same construct as when it is defined by linguists and non-experts, or whether the meaning of coherence is different for attachment experts who apply it to attachment interviews. If there is no difference in the definition and use of coherence among the various groups, attachment interviews might be coded with a measure for coherence by other coders than attachment experts, or even with the help of advanced computer programs.

It may not be necessary to have thorough knowledge of attachment theory and research to be able to observe coherence in attachment interviews. Knowledge of linguistics may be sufficient since attachment interviews are coded on the basis of properties of the discourse, which are consonant with Grice's maxims. Moreover, the question is whether training in Grice's maxims is necessary. Grice suggested that rational language use presupposes adhering to the four maxims. So it may even be that competent language users, without any education in attachment theory or Grice's maxims, are intuitively able to define coherence adequately. In contrast, it might be argued that competent language users may not be able to make their underlying assumptions explicit when defining coherence in attachment interviews. To get insight into this matter we have conducted an empirical conceptual study of coherence.

In the present study, the ideas of four groups of participants regarding coherence were compared: attachment experts, linguists, higher educated non-experts, and lower educated non-experts. In a preliminary analysis, we investigated the associations among the four maxims (quality, quantity, relation, and manner) of coherence. Since they represent different concepts, they were not expected to be strongly associated. We then examined whether we could correctly predict if participants were attachment experts or not. It was hypothesized that we would be able to distinguish the attachment experts from the combined other groups because of the difference in specific attachment-related expertise. Concerning the separate groups, we expected that it would be difficult to predict whether participants were attachment experts or linguists, because both have been educated in Grice's maxims. We hypothesized that we would be able to distinguish attachment experts from the two groups of non-experts, because of the difference in education regarding coherence. More specifically, it was expected that attachment experts would mainly differ from non-experts in emphasizing the maxim of quality. The maxim of quality is suggested by Grice as the most important maxim, so we expected that attachment

experts would emphasize this maxim more heavily than non-experts would do. Including a group of lower educated non-experts enabled us to test whether a certain level of education is a necessary or sufficient condition to describe coherence similarly to attachment experts. When, contrary to our expectations, both higher and lower educated non-experts cannot be distinguished from the attachment experts, we may conclude that every rational language user should be able to rate coherence in attachment interviews.

Method

Participants

Thirty two participants were involved in the present study: 9 experts in the field of attachment, 6 linguists, 8 higher educated (HE) non-experts, and 9 lower educated (LE) non-experts. The selection criterion for attachment experts was participation in an AAI Institute. This Institute is an intensive, two week training workshop in coding the AAI. During the training, attachment and coherence are central constructs. The linguists were required to have obtained a PhD in their field. The two groups of non-experts were required to lack any specific knowledge of attachment theory or linguistics. The educational level of the LE non-experts was senior secondary vocational education or higher vocational education and the HE non-experts' education was at PhD-level. Table 1 presents an overview of the background variables for all groups.

The attachment experts were all female and between 29 and 67 years old (M = 46.9 years, SD = 10.8). Six had the Dutch nationality, two were American and one was British. They had obtained a Master degree or PhD (M = 6.9, SD = 0.3, on a scale ranging from 1 = pre-school to 7 = post-doctorate/PhD) and had participated in an AAI Institute between 1995 and 2002. The linguists were male, aged 53 to 66 years (M = 57.3 years, SD = 5.0). Four of them were Dutch, one was Dutch/New Zealander and one had the Danish nationality. As indicated, they all had a PhD in linguistics. The group of higher educated non-experts consisted of eight women between 33 and 54 years of age (M = 41.6 years, SD = 8.1). All of them had the Dutch nationality, except for one, who was from New Zealand. They had completed a PhD. The group of lower educated non-experts consisted of two males and seven females. They were between 23 and 71 years old (M = 38.0, SD = 14.4). Eight of them were Dutch and one had the British nationality. The mean educational level of the LE non-experts was 4.4 (SD = 0.5).

Table 1

Background Variables of the Participants

Variable	Attachm experts	Linguists	HE non-experts	LE non-experts	Total
N	9	6	8	9	32
Female (%)	100.0	0.0	100.0	77.8	75.0
Dutch nationality (%)	66.7	83.3	87.5	88.8	81.3
Age (years)	46.9 (10.8)	57.3 (5.0)	41.6 (8.1)	38.0 (14.4)	45.0 (12.2)
Educational level[a]	6.9 (0.3)	7.0 (0.0)	7.0 (0.0)	4.4 (0.5)	6.3 (1.2)

Note. Attachm = Attachment. HE = higher educated. LE = lower educated. Standard deviations are shown in parenthesis.

[a]Educational level is assessed with a scale ranging from 1 = pre-school to 7 = post-doctorate/PhD.

The four groups of participants differed significantly with regard to age (F (3, 28) = 4.37, p = .01), gender (χ^2 = (3, N = 32) 22.96, p < .01), and educational level (χ^2 (3, N = 32) = 28.24, p < .01). The significant difference in educational level follows from the inclusion of lower educated non-experts in the sample.

Measures

Coherence Q-sort (CQS)

The Coherence Q-sort was developed as a measure for assessing coherence of discourse in interviews about attachment.[3] The instrument is based on the guidelines for scoring and classifying the AAI (Main et al., 2003). Four experts, who were trained in the AAI scoring and classification systems, each independently formulated 30 items indicating either coherence or incoherence. From this set of 120 items, 78 items were used in a pilot study of 32 interviews. During the pilot, items were adjusted when necessary and discarded when redundant. In addition, we added filler-items and items concerning the way in which the interview had been conducted. In its final state, the Coherence Q-sort consists of 72 items (see Table 4). Seven of these items are fillers and three items concern the interview protocol, the interview context, or the performance of the interviewer. The other 62 items pertain to coherence. More specifically, they refer to one of Grice's four maxims, as follows: 22 items concern quality, 10 items concern quantity, 8 items focus on relevance, and 22 items focus on manner (see Table 2). Examples of items indicative of the maxim of quality are "has evidence for what he says" and, as indicative of a violation of the maxim of quality, "contradicts himself during the interview without noticing". An item concerning quantity

[3] The developers of the CQS and the participants in this study are different persons.

is “answers in an extremely concise way”. Items indicative of relevance and manner are respectively “does not drift away from the main topic of the question” and “does not substitute nonsense words for parts of the sentences”.

The 72 items of the CQS are sorted into nine piles, ranging from *does not fit at all* with the interview to *fits very well* with the interview. The distribution of the items is forced and uniform, with eight items per pile.

Procedure

The participants were asked to sort the Coherence Q-sort for the hypothetically most coherent interview transcript (below referred to as ‘prototypical coherent interview’) such as the Adult Attachment Interview. Rather than giving the participants an interview transcript, we asked them to imagine what the ideally coherent interview would look like. They were informed that in these interviews participants are asked for general evaluations of relationships and/or events as well as concrete evidence supporting these evaluations. Furthermore, we instructed them in Grice’s maxims and gave descriptive illustrations of violations of these maxims. Finally, they were asked to put the three items about the interview protocol, the interview context, and the interviewer in the middle pile, because these are not applicable when sorting the CQS for a prototypical coherent interview. Background information of the participants, such as gender, age, and educational level was obtained with a short questionnaire.

Reliability

As can be seen from Table 2, interrater reliabilities for coherence ranged from .67 for the lower educated non-experts to .86 for the attachment experts. The reliabilities for Grice’s maxims were satisfactory for all groups of sorters (see Table 2). For coherence as well as for the separate maxims, the reliabilities for the combined groups were also adequate.

Data-analysis

Items indicative of incoherence were recoded into reverse order, and scores for Grice’s maxims were calculated as the average score of the corresponding items. First, we calculated the correlations among the maxims. Second, means and standard deviations for each maxim were computed per group. For each group of participants ANOVAs and Kruskal-Wallis tests were conducted to examine whether the maxims differed significantly from each other. Finally, discriminant analyses were conducted to predict group membership from Grice’s criteria.

Table 2

Interrater Reliabilities of the CQS

		Interrater reliability			
Category	Items (N)	Attachm experts	Linguists	HE non-experts	LE non-experts
Quality	22	.87	.84	.64	.72
Quantity	10	.81	.86	.71	.67
Relevance	8	.88	.94	.75	.88
Manner	22	.85	.75	.63	.59
Coherence[a]	62	.86	.84	.67	.79

Note. The seven filler-items and the three interview items were omitted. Attachm = attachment. HE = higher educated. LE = lower educated.

[a]The Coherence scale consists of the items of the 4 scales (Quality, Quantity, Relevance, and Manner).

Results

Correlations

Table 3 shows the correlations among the four maxims. Relevance and manner were negatively correlated: participants who assigned higher values to manner, considered relevance of less importance.

Differences on Grice's maxims

Means and standard deviations per item are presented in Table 4, and means and standard deviations for each maxim are shown in Table 5. For each of the four groups of participants, ANOVAs and Kruskal-Wallis tests were conducted to examine whether the maxims were valued differently. Both parametric and non-parametric tests were significant for all groups (see Table 5 for parametric statistics[4]). Tukey post hoc tests revealed that attachment experts valued quality and relevance significantly more than quantity. Linguists and both groups of non-experts emphasized relevance more than the other three maxims. Finally, linguists gave more weight to quantity than to manner.

[4] Statistics of the non-parametric tests were similar to those of the parametric tests.

Table 3

Correlations between Grice's Maxims

	Quality	Quantity	Relevance	Manner
Quality	--	-.14	-.13	.24
Quantity		--	.26	-.25
Relevance			--	-.42*
Manner				--

Note. $N = 32$.

*$p < .05$.

Predicting group membership

Discriminant function analysis (DA) was performed using Grice's maxims as predictors of membership of two groups: attachment experts versus the combined other groups. The latter group consisted of linguists, higher educated non-experts and lower educated non-experts. In addition, with an exploratory aim, we conducted three discriminant analyses to distinguish the attachment experts from each of the other three groups separately. Although sample sizes of the groups were small (range: 6 to 23), DA could be performed because the sample size of the smallest group still exceeded the number of predictor variables (Tabachnick & Fidell, 2001). Evaluations of the assumptions of DA revealed no serious threat to multivariate analysis.

Table 4

Content Categories, Group Means, and Standard Deviations per Item

Item	Item description	Category	Attachm experts	Linguists	HE non-experts	LE non-experts
54	is consistent, that is, later information is consistent with earlier information	Qual	8.9 (0.3)	8.8 (0.4)	8.6 (0.7)	8.4 (0.7)
41	has evidence for what he says	Qual	8.8 (0.4)	8.5 (0.8)	8.6 (0.5)	8.0 (0.9)
61	uses fresh, authentic speech	Man	8.8 (0.7)	7.5 (1.2)	6.8 (1.2)	7.2 (1.5)
7	displays metacognitive monitoring, that is, reflects on the processes of thinking and recall that take place during the interview. For example '…Oh dear, that's completely contradictory to what I just said.'[a]	Qual	8.6 (1.0)	7.2 (1.0)	6.6 (1.8)	5.2 (2.5)
1	provides sufficient context for the interviewer to be able to understand the answers	Quan	8.3 (0.9)	8.8 (0.4)	8.6 (0.7)	7.8 (0.7)
42	reasonable evaluation of effects of experiences or events on self	Qual	8.3 (0.7)	7.3 (0.8)	7.8 (0.7)	7.7 (0.9)
24	provides adequate illustrations with general evaluations when asked for	Qual	8.2 (0.8)	8.7 (0.5)	8.4 (0.7)	7.8 (0.8)
20	is cooperative, for example keeps the interviewer informed about his reasoning	Qual	8.0 (1.1)	8.7 (0.5)	7.5 (0.9)	8.6 (0.7)
72	responds consistently, but with varied answers throughout the interview[b]	Qual	8.0 (1.1)	8.0 (0.9)	8.5 (0.8)	5.2 (3.2)
47	is involved in the interview without losing track of the interviewer	Quan	7.9 (1.1)	8.3 (0.8)	7.5 (1.2)	8.4 (0.7)
6	does not drift away from the main topic of the question	Rel	7.7 (0.5)	8.7 (0.5)	8.6 (1.1)	8.0 (0.7)
55	does not avoid answering a question by addressing another issue	Rel	7.7 (0.9)	7.8 (1.2)	7.9 (1.2)	6.7 (2.5)
68	the interview can easily be understood	Man	7.7 (1.1)	7.7 (1.5)	7.0 (1.8)	7.8 (1.1)

Item	Item description	Category	Attachm experts	Linguists	HE non-experts	LE non-experts
38	shows that he is still reflecting on some of his opinions	Qual	7.6 (0.9)	6.3 (0.5)	7.0 (1.6)	6.4 (2.5)
15	does not finish discussion of topic prematurely	Quan	7.3 (0.5)	8.7 (0.5)	8.0 (0.9)	7.4 (1.6)
17	talks about relationships in a spontaneous way	Man	7.3 (1.2)	6.2 (1.0)	7.1 (1.0)	7.0 (1.5)
40	presents a personal account of events	Man	7.3 (0.9)	6.8 (0.8)	6.9 (1.6)	7.2 (2.0)
65	does not substitute nonsense words for parts of the sentences	Man	7.3 (0.7)	6.3 (2.2)	6.6 (0.7)	7.2 (0.8)
9	no remarkable slips of the tongue, for example no slips of the tongue in which the person talked about is confused with the self[c]	Man	7.2 (1.1)	6.5 (0.8)	6.4 (0.9)	6.8 (1.4)
56	does not use exaggerated language in the context of the discourse	Man	7.2 (1.1)	6.8 (1.2)	7.3 (0.5)	7.2 (1.0)
59	does not trail off, finishes sentences	Man	7.2 (0.7)	7.2 (1.0)	7.5 (0.9)	8.2 (0.7)
63	reader agrees with evaluations of events and persons	Qual	7.2 (1.5)	5.5 (1.4)	5.5 (1.3)	6.2 (1.4)
2	varied description of (different) people, variation in thematic content of adjectives	Qual	7.0 (0.9)	6.5 (1.4)	6.0 (1.2)	7.0 (1.3)
57	corrects the interviewer when he incorrectly mirrors his answers	Qual	7.0 (1.3)	8.2 (1.2)	8.4 (0.7)	7.4 (2.2)
4	does not present exaggerated comparisons, as in ‘my childhood was like I was in prison’	Man	6.8 (1.0)	6.8 (0.4)	6.3 (1.4)	4.7 (2.3)
45	thinks before answering	Man	6.6 (1.0)	7.5 (0.5)	8.5 (0.5)	8.6 (0.7)
66	talks about ‘I’, not about ‘you’ and ‘we’	Man	6.6 (0.9)	5.2 (1.0)	6.6 (1.7)	5.4 (2.6)
14	uses colorful language	Fill	5.8 (0.4)	6.2 (0.8)	5.5 (1.1)	6.1 (1.6)
16	uses grammatically correct sentences	Fill	5.8 (0.7)	6.0 (1.1)	6.9 (1.0)	7.3 (1.3)
28	is humorous	Fill	5.8 (0.8)	6.5 (1.5)	5.8 (1.5)	6.1 (0.8)
27	frequently asks the interviewer to clarify questions	Fill	5.3 (0.9)	5.5 (1.9)	6.5 (1.3)	5.1 (1.9)
22	uses slang	Fill	5.1 (0.6)	4.2 (1.5)	4.0 (1.3)	3.2 (1.7)
53	frequently uses the word ‘however’	Fill	5.1 (0.6)	5.5 (0.8)	5.1 (1.1)	4.3 (1.2)
23	provides a factual story without emotional evaluations	Man	4.9 (0.8)	6.7 (1.0)	6.4 (1.2)	6.3 (2.1)

Item	Item description	Category	Attachm experts	Linguists	HE non-experts	LE non-experts
60	provides more information than needed	Quan	4.8 (0.8)	2.5 (1.0)	4.1 (2.0)	5.1 (1.5)
46	uses rude language	Fill	4.7 (1.4)	4.3 (1.0)	2.8 (1.3)	2.6 (1.7)
34	tells complex or long story	Quan	4.6 (0.9)	2.8 (0.8)	3.1 (1.8)	2.4 (1.1)
13	describes limited repertoire of events	Qual	4.0 (1.0)	4.2 (1.5)	3.9 (1.2)	5.7 (1.5)
35	is reluctant to speak	Quan	3.9 (1.4)	3.3 (1.2)	2.4 (1.1)	3.4 (1.7)
52	answers in an extremely concise way	Quan	3.7 (1.9)	5.2 (2.6)	7.1 (2.2)	5.3 (2.1)
18	fails to provide general description	Qual	3.6 (1.3)	3.8 (1.2)	3.6 (0.7)	3.2 (1.9)
21	discusses issues that are not relevant for the interview	Rel	3.4 (1.0)	1.8 (0.8)	2.4 (1.4)	2.6 (1.2)
49	often uses general statements, for example ‘my mother always took me to the swimming pool’ or ‘my mother is a nice person’	Qual	3.4 (0.7)	4.3 (1.4)	4.4 (1.3)	4.2 (1.0)
32	talks about persons or events that were not introduced	Man	3.0 (1.7)	2.5 (0.8)	2.9 (1.6)	4.0 (1.9)
37	uses long (not broken) sentences containing several different messages	Man	2.9 (1.1)	3.0 (0.9)	3.5 (1.9)	3.2 (1.9)
58	fails to recall events	Qual	2.9 (1.3)	2.5 (1.0)	2.8 (1.0)	2.6 (1.4)
62	frequently uses stopgaps or fillers rather than finishing (part of) a sentence, for example ‘… and this and that.’[d]	Man	2.9 (1.3)	3.2 (1.0)	3.3 (1.2)	3.3 (1.5)
69	upbeat endings to descriptions of negative aspects/events	Qual	2.8 (2.0)	5.7 (0.8)	3.4 (1.2)	3.4 (1.5)
8	tries to convince the interviewer	Man	2.7 (1.2)	4.7 (0.8)	5.1 (1.4)	5.4 (1.5)
11	often uses global evaluations without concrete illustrations	Qual	2.7 (1.0)	3.0 (0.9)	3.6 (1.1)	3.6 (1.7)
36	uses angry or accusative language	Man	2.7 (1.2)	3.2 (1.5)	2.8 (1.0)	2.1 (1.5)
70	repeats himself in answering different questions, without adding new information	Quan	2.7 (0.9)	1.8 (1.0)	3.3 (2.1)	3.2 (1.9)
48	is incomplete, that is, provides insufficient information	Quan	2.6 (1.0)	1.8 (1.0)	2.1 (1.6)	3.0 (1.2)

Item	Item description	Category	Attachm experts	Linguists	HE non-experts	LE non-experts
51	uses canned speech such as jargon, slogans and psychobabble	Man	2.6 (1.1)	3.3 (1.4)	2.8 (1.5)	3.4 (1.6)
5	goes on and on about a subject	Quan	2.4 (0.9)	2.2 (0.8)	2.5 (0.9)	2.9 (0.9)
44	answers as if he did not hear the question	Rel	2.4 (1.0)	1.7 (1.2)	1.6 (0.7)	2.0 (1.1)
50	often quotes (other) persons without introducing the quotation properly	Man	2.4 (1.2)	3.2 (1.2)	2.1 (0.6)	4.0 (1.8)
25	often forgets the question that started his response	Rel	2.3 (0.9)	1.8 (0.8)	2.0 (1.1)	2.4 (1.7)
30	circles around the subject, that is, does not come to the point	Rel	2.3 (1.2)	1.3 (0.5)	1.5 (0.8)	2.1 (0.8)
3	supports general evaluations with (other) general evaluations rather than with concrete illustration	Qual	2.2 (1.1)	3.8 (2.0)	4.0 (1.7)	3.8 (1.6)
10	uses non-logical explanations or interpretations	Qual	2.2 (1.1)	2.2 (1.0)	1.9 (1.1)	2.3 (1.0)
39	confuses past and present tense in talking about the past several times	Man	2.2 (1.1)	4.3 (0.8)	3.0 (1.4)	3.3 (1.5)
26	contradicts himself during the interview without noticing	Qual	2.0 (1.0)	1.2 (0.4)	1.8 (1.2)	2.0 (1.2)
31	changes evaluations within the interview without noticing	Qual	2.0 (0.9)	1.7 (1.2)	2.6 (1.4)	2.8 (1.3)
64	repeatedly brings up the same not-relevant issue	Rel	2.0 (0.9)	1.2 (0.4)	2.3 (1.3)	2.8 (1.5)
43	displays ‘mindreading’, that is, thinks that he completely knows other people’s motives	Qual	1.8 (0.8)	3.7 (1.0)	2.8 (1.3)	3.0 (1.1)
19	gets lost in the interview, loses sight of the interview context	Rel	1.7 (0.7)	1.3 (0.5)	1.6 (0.7)	1.1 (0.3)
71	is angry with the interviewer as if he is part of the story	Man	1.4 (1.0)	2.3 (1.5)	2.4 (0.9)	2.4 (1.7)
67	uses convoluted, entangled sentences that remain difficult to comprehend even after re-reading	Man	1.1 (0.3)	2.0 (0.9)	1.9 (1.1)	1.8 (0.7)

Item	Item description	Category	Attachm experts	Linguists	HE non-experts	LE non-experts
12	the interviewer conducts the interview appropriately, for example probes when necessary and does not ask suggestive questions	Interv				
29	the interview is conducted in an appropriate context, for example there are no interruptions by persons or a telephone	Interv				
33	the interview protocol consists of general questions to elicit a general description, and specific questions to elicit evidence for the description	Interv				

Note. Attachm = Attachment. HE = higher educated. LE = lower educated. Qual = Quality. Quan = Quantity. Rel = Relevance. Man = Manner. Fill = Filler. Interv = Interview item. Participants were instructed to put interview items in pile 5 since there was no actual transcript for which the items were sorted; therefore, no means are presented for these items.

[a]recognition of an appearance-reality distinction; recognition of representational diversity; recognition of representational change; monitoring for, for example possible logical contradictions.

[b]consistent & varied -> item is placed high, consistent but not varied -> lower, not consistent but varied -> even lower, not consistent & not varied -> lowest.

[c]an example of a slip is 'I died when my father was 14'.

[d]'You know' does not fall into this category.

Table 5

Means and Standard Deviations for Grice's Maxims per Group

	Quality		Quantity		Relevance		Manner				
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*F*	*df*	*Post hoc*
Attachment experts	7.64	0.43	6.90	0.28	7.64	0.41	7.33	0.36	7.94**	3, 32	rel qual > quan**
Linguists	7.17	0.37	7.62	0.34	8.42	0.25	6.64	0.38	29.48**	3, 20	rel > qual man*** rel > quan** quan > man***
HE non-experts	7.19	0.37	6.95	0.52	8.14	0.39	6.86	0.28	17.51**	3, 28	rel > qual quan man**
LE non-experts	6.88	0.45	6.76	0.45	7.71	0.51	6.72	0.50	8.55**	3, 32	rel > qual quan man***
Others combined	7.07	0.42	7.05	0.56	8.04	0.49	6.75	0.39	33.18***	3,88	rel > qual quan man***

Note. HE = higher educated. LE = lower educated. Qual = Quality. Quan = Quantity. Rel = Relevance. Man = Manner.

**$p < .01$. ***$p < .001$.

Attachment experts versus combined other groups

When predicting group membership for attachment experts versus the combined other groups a significant function was found, χ^2 (4, N = 32) = 20.58, p < .01. Three of Grice's maxims significantly distinguished the attachment experts from the combined other groups (quality, relevance, and manner, see Table 6). Quality (.61) and manner (.67) were the best predictors; relevance had a somewhat lower loading (-.38) on the discriminant function. Attachment experts ranked quality and manner items higher than the combined other groups, whereas the combined other groups assigned more weight to relevance than attachment experts. With the jackknifed classification procedure, 27 participants (84.4%) were classified correctly. There were no attachment experts among the incorrectly classified participants.

Attachment experts versus separate groups

Discriminant analyses were also performed for attachment experts versus linguists, attachment experts versus HE non-experts, and attachment experts versus LE non-experts, to explore the specific differences between these groups. First, the attachment experts could be reliably separated from the linguists, χ^2 (4, n = 15) = 23.36, p < .01. All four maxims were significant predictors. Attachment experts scored higher on quality and manner than linguists, whereas linguists emphasized quantity and relevance more than attachment experts. When classifying these participants in one of the two groups, all but one (n = 14, 93.3%) were correctly classified. The participant who was classified incorrectly was a linguist.

When distinguishing attachment experts from higher educated non-experts, a significant function was found (χ^2 (4, n = 17) = 16.19, p < .01) in which quality, relevance, and manner were significant predictors. Attachment experts had higher scores on quality and manner than the HE non-experts, whereas HE non-experts assigned more weight to relevance than attachment experts. The number of correctly classified participants was also high for this function: 14 out of 17 (82.4%). One attachment expert and two HE non-experts were incorrectly classified.

Finally, the attachment experts were reliably separated from the lower educated non-experts, χ^2 (4, n = 18) = 16.28, p < .01. Attachment experts emphasized quality and manner significantly more than the LE non-experts. Of the 18 participants, 17 (94.4%) were correctly classified. The only incorrectly classified participant was a LE non-expert.

Table 6

Discriminant Function Analyses of Grice's Maxims

	Attachment - Others combined		Attachment - Linguists		Attachment - HE non-experts		Attachment - LE non-experts	
Predictor	Corr	*F* (1,30)	Corr	*F* (1,13)	Corr	*F* (1,15)	Corr	*F* (1,16)
Qual	.61	12.06**	.22	4.81*	.37	5.16*	.61	13.21**
Quan	-.13	0.57	-.46	19.89**	-.04	0.06	.14	.66
Rel	-.38	4.77*	-.43	17.39**	-.43	6.74*	-.05	.10
Man	.67	14.80**	.36	12.65**	.49	8.99**	.50	8.84**

Note. Attachment = Attachment experts. Others combined = Combined other groups. HE = higher educated. LE = lower educated. Qual = Quality. Quan = Quantity. Rel = Relevance. Man = Manner. Corr = Correlations of predictors with the discriminant function.

*$p < .05$. **$p < .01$.

Discussion

The present study demonstrates that attachment experts can be reliably distinguished from the (combined) other groups with regard to defining coherence in attachment interviews. When examining the differences between attachment experts and each of the three other groups separately, we found that (1) attachment experts had higher scores on quality and manner than linguists, whereas linguists emphasized quantity and relevance more; (2) attachment experts assigned also more weight to quality and manner than higher educated non-experts, whereas higher educated non-experts had higher scores on relevance; (3) again, attachment experts emphasized quality and manner more than lower educated non-experts. Within group comparisons showed that attachment experts valued quality and relevance more than quantity, while all other groups gave more weight to relevance than to any of the other maxims. In contrast to our hypothesis, not only the HE and LE non-experts but also the linguists could be reliably distinguished from attachment experts. It seems that knowledge of Grice's maxims is not enough to define coherence in attachment interviews similarly to attachment experts; knowledge of attachment theory appears to set them apart from linguistic experts.

How can we explain the differences we found between the attachment experts and the other three groups? A possible reason for the stronger emphasis of attachment experts on quality may be that they have followed Grice (1975) more in his suggestion that quality can be seen as the most important of the four maxims.[5] In the Main et al. (2003) coding system quality also is heavily emphasized. In the field of linguistics, some researchers only distinguish relevance and quantity, whereas others consider relevance as the uniquely significant maxim. Relevance and quantity have therefore received much attention (see Haberland & Mey, 2002, for a review), which may have led linguists to emphasize quantity and relevance. Attachment experts scored also higher on quality and manner than both groups of non-experts. These maxims are possibly the most difficult to comprehend for participants who are new to this theory, and the non-experts may therefore have given less weight to these two maxims. The results of our study are in agreement with the findings of a study on computers imitating human language use (Saygin & Cicekli, 2002). Linguists and non-experts may have seen violations of manner as human-like and have therefore given less weight to this maxim, while stressing the maxims of quantity and relevance, which were indeed marked as violations uncovering the identity of the computer in Saygin and Cicekli's (2002) study.

[5] Note that as a consequence of the use of the q-sort methodology a group of participants cannot have higher scores on all maxims than any other group.

The different theoretical background of attachment experts and linguists might also explain the differences in their ratings of dimensions of coherence. In the AAI Grice's maxims are used not merely to assess coherence but also for the purpose of classifying adults as demonstrating a secure or insecure attachment representation. An important difference with linguists and non-experts is that in attachment interviews, attachment experts as coders try to infer the underlying defensive mechanisms used by the respondent to process attachment-related experiences (George & West, 2001; Main, Kaplan, & Cassidy, 1985). These processes operate mostly outside awareness (Bowlby, 1980; Main, 1990, 1999). Dismissing adults, for example, would typically show deactivation of the attachment system (Cassidy & Kobak, 1988; George & West, 2001; Solomon, George, & De Jong, 1995). These adults systematically exclude attachment-relevant information from the consciousness to minimize mental suffering (Bowlby, 1980; Main, 1999; Peterfreund, 1971). As a result, dismissing respondents are typically unable to give evidence for what they claim was a perfectly normal or very nice childhood. In these cases attachment experts may be likely to judge the interview as more incoherent than linguists, who may interpret this lack of evidence as common lack of memory. In contrast to what is the case for linguists, the concept of coherence might thus refer to an underlying psychological component when assessed by attachment experts in attachment interviews. This explanation can easily be tested, by systematically varying the type of interview that is considered, extending the range to non-attachment, non-psychological interviews.

The results of the current study have important implications for the application of q-sorts for coherence in attachment research. Coherence as measured by the CQS is not defined similarly by attachment experts and linguists or non-experts. It seems that the CQS cannot be applied to attachment interviews without training in attachment theory and research. Our study does not provide evidence that the CQS has the potential for making the scoring of AAIs more accessible for non-attachment experts. Future research should investigate whether psychologists and clinicians without specific training in the AAI define coherence similarly as attachment experts. When they do, there would be the opportunity to make the coding of attachment interviews more accessible for non-attachment experts by using the CQS. It is, however, an empirical question whether they are able to apply their psychological knowledge to the assessment of coherence in attachment interviews without specific training in coding attachment interviews.

Applying the concept of coherence to attachment interviews requires more than linguistic knowledge. It can therefore be concluded that the requirements for a computer program capable of coding AAIs are far beyond the current state of the art in computerized text analysis. Counting words that are indicative of expressed affect or measuring emotion-abstraction patterns only partially overlaps with analyzing

discourse characteristics and their psychological meaning. The possible underlying psychological component of coherence might be uncovered by asking attachment experts to think aloud (Ericsson & Simon, 1993) when coding AAIs. Vital rules might then be derived that can be imputed into computer programs. However, connecting general statements in the beginning of the interview with more specific evidence in later parts requires narrative competence still beyond the power of the computer. Coding AAIs will therefore remain an activity that, unfortunately, takes a lot of human effort.

Although the CQS may not be an easy alternative to the Adult Attachment scoring and classification systems (Main et al., 2003), an important aspect of the CQS might be implemented in the classification systems. The CQS distinguishes various aspects of coherence, whereas in the traditional coding system indications of coherence and incoherence are only summarized in one score on a 9-point rating scale. Additional 9-point rating scales may be developed for each of the four maxims so that the different aspects of (in)coherence can be assessed and examined. Such rating scales would make it possible to test whether violations of specific maxims are associated with specific types of parental insensitivity. Our study makes clear that the various maxims may indeed index different dimensions and these may be related to different behavioral implications.

This study also contributes to the discussion surrounding Grice's maxims as described in the Introduction. When inspecting the correlations between the maxims (range: -.42 to .26), it appears that coherence cannot be characterized by only one dimension. If the categorization of the items was somehow inaccurate, interrater reliabilities would have been lower and correlations higher. Therefore, the empirical evidence suggests that Grice's maxims cannot conceptually be reduced to one dimension. The unexpected negative association between manner and relevance might be a consequence of a choice participants made between content and form. The maxim of manner refers to the clarity of the words and sentences being used, while the focus of relevance is whether the content of an answer is in agreement with what the person is asked for.

In the present study the participants did not sort the CQS for an actual Adult Attachment Interview transcript. We asked the sorters what in their opinion were the characteristics of the ideally coherent interview, and this may not be identical to observing which maxims they actually would rely on when assessing the coherence of an interview transcript. This may be seen as a limitation of our study. However, differences among coders in the interpretation of (parts of) a real interview transcript would confound their scores with their definition of what is essential for coherence.

When interpreting the results it should be noted that this is a pilot study: an exploratory conceptual study on coherence in a relatively small sample. Future studies are needed to draw more definite conclusions. Nevertheless, our various analytical strategies point into the same direction: Defining coherence in attachment interviews is more than just applying Grice's linguistic maxims; expertise in attachment theory is determining the way in which coherence of attachment interviews is defined.