

Measuring the complexity of viewers' television news interpretation: Differentiation

GABI SCHAAP, RUBEN KONIG, KARSTEN RENCKSTORF
and FRED WESTER

Abstract

If television news viewers are conceived as active audience members, their interpretations should be a crucial factor in the study of the 'effects' of television news. Here, viewers' interpretations are understood as subjective (re)constructions of a news item. In a previous contribution, we argued that interpretations can vary both within and between viewers in regard to the level of complexity. Complexity is the degree to which interpretations are a) differentiated, and b) integrated. In this contribution, we will operationalize the concept of differentiation of television news interpretations by its viewers. Furthermore, we will present a procedure for measuring differentiation based on the thoughts viewers reported while they watched a television news program. Results of a small-scale study (N = 19) provided first indications that the procedure is able to discriminate between viewers with varying levels of differentiation in interpreting television news.

Keywords: television news, viewers' interpretation, interpretive complexity, differentiation, integration, Thought-Listing Technique

Interpreting television news is a complex process in which viewers (re-)construct a news item into something that has a meaning and that makes sense to them. If news viewers are to be seen in this fashion, as actively reshaping the content of news, this should have consequences for the way the impact of television news is studied. Despite increasing consensus on this matter, the way in which news interpretation should be conceptualized and measured remains unclear, and as a consequence empirical studies on the subject have been scarce (Gunter, 2001; Livingstone, 1989; Renckstorf and Wester, 2001). In an attempt to contribute to this research area, we recently introduced the concept of interpretive complexity, arguing that studying the structural components of interpretations may be a useful addition to the field (Schaap, Renckstorf, and

Wester, 2005). In this article we focus on operationalizing and measuring one aspect of this concept, called differentiation. Differentiation refers to the specificity and heterogeneity of interpretations (cf. Schaap et al., 2005).

Together, the previously introduced concept and the operationalizations and measurement procedure presented here, have the potential of yielding a more direct and highly detailed image of how people interpret the news. This contribution should not so much be read as a report of empirical findings, but as a proposal for a method to study television news interpretation. We use a small-scale study to illustrate the method and test the usefulness of describing interpretation in terms of structural characteristics by assessing variations in the complexity of different television news viewers' interpretations.

The structure of viewers' interpretations: Differentiation and integration

Viewers make sense of television news by constructing a representation, or interpretation, of a news item (cf. Renckstorf and Wester, 2001). This representation is not a direct copy of the program. Instead, we can look at it as made up of subjective interpretations, formed by personal and social knowledge as well as the news. According to our concept of 'interpretive complexity', all interpretations of the news share at least two general structural characteristics (Schaap et al., 2005). That is, viewers make sense of the news by a) using basic elements from the news and/or from their own knowledge, and by b) connecting these elements to create a subjective, meaningful whole.

The degree to which a viewer's interpretation contains large or small amounts and a broad or narrow range of basic elements is called interpretive differentiation, which is the subject of this study. Inspired by the work of James Spradley, we argue that a viewer constructs a representation of a television news item that may incorporate references to specific actors, with goals and feelings, acts, to events and activities, and objects, and to time and space (Spradley, 1972, 1979, 1980; Spradley and McCurdy, 1972). These actors, objects, etc. have specific attributes; they have causes, reasons, functions, etc. The first characteristic of differentiation is the number of single specific occurrences of these elements in the interpretation. The larger the number of specific actors, acts, attributes, causes, etc. incorporated in an interpretation, the more 'specific' it is (Schaap et al., 2005).

In addition to the number of elements, we should take into account the range, or heterogeneity of elements. One viewer's interpretation of a news item may contain, for instance, several actors, such as Jacques Chirac, Gerhard Schröder, and George W. Bush. An interpretation by

another viewer may contain just one actor, for instance George W. Bush, and in addition an act (e.g., voting), and an object (e.g., an amendment). Both of our exemplar viewers use three elements. However, the first viewer uses three elements of the same type; all of them actors. The second viewer uses three elements of three different types: an actor, an act, and an object. Therefore, although the number of elements used by both viewers is equal (i.e., both interpretations are equally specific), the range of elements is different (the second interpretation is more heterogeneous than the first). Accordingly, interpretive differentiation has two aspects: the number of specific elements and the number of types of elements (or: range). An interpretation containing many elements and many types of elements is called more differentiated than an interpretation containing fewer elements of less different types.

In the interpretation of a television news item, the many (or not so many) single elements are connected to other elements. Representing the coherence of an interpretation, this second structural characteristic of interpretive complexity is called integration. The measurement of integration will be discussed elsewhere, so it will suffice to say that the integration of an interpretation is determined by the explicit, causal, logical, and temporal connections that are made between individual elements, as well as by the grouping of elements into larger socio-cultural categories (cf. Schaap et al., 2005).

As different viewers apply different personal and social knowledge to construct meaning, we can assume that the interpretation of a given news item varies between different viewers. Some viewers may have a highly differentiated (and/or integrated) interpretation of a news item, whereas other viewers have a much less differentiated (and/or integrated) interpretation of the same item. Furthermore, a given viewer may have a differently structured interpretation of different news items. That is, the same viewer may interpret one item in a much more complex way than another item. In the following we present a way to measure differentiation as defined above.

Measuring interpretive differentiation

The method for measuring viewers' interpretations of television news consists of four components. First, a data gathering instrument to 'tap' viewer's thoughts the moment they are watching the news, and second, a three-step procedure to assess differentiation in reported thoughts.

Data gathering: Thought-Listing Technique

In order to allow the participants to communicate their interpretations freely and directly, we adapted and tested an instrument called Thought-

Table 1. *News bulletin for Thought-Listing Technique, NOS 8 o'clock news, 21-11-2000.*

Item	Issue	Description	Length
1	Profession-related diseases	Company doctors fail to report sick employees	3:18
2	BSE	The Netherlands will be testing cattle earlier and more often	2:54
3	Israel	Egypt withdraws its ambassador from Israel after rocket attacks on Palestine territories	2:36
4	Euthanasia	Euthanasia directive used by family members to manipulate physicians into euthanasia when care proves too difficult	2:30
5	Exhibition	Queen Beatrix and president Rau open exhibition on Dutch-German relations	2:42
6	Emmy Awards	TV series 'All Stars' wins American Emmy award in 'best drama series' category	2:24
7	Weather forecast		1:18

Note. Item labels are ours; Item 1 was used as a practice item and was excluded from the analyses, as was the weather forecast

Listing Technique (Schaap, 2004). This observation instrument enables the participants to report all thoughts at the moment they occur while watching a news program.

We showed an eighteen-minute videotaped news bulletin, containing seven items, to nineteen participants (Table 1). The participants were selected to reflect a wide variety in terms of sex, age, and education¹. Additionally, the broadcast was edited so that the screen turned 'black' after small 'natural' segments of the news, i. e., so as not to disrupt the normal flow of a news item too much. The segments averaged eighteen seconds in length. Participants were asked to say out loud all thoughts they had while they were watching the news segment at each interval (for a more detailed description of the procedure and its logic, cf. Schaap, 2004). These verbalizations were recorded and subsequently transcribed, resulting in nineteen protocols of verbalized thoughts, with an average length of about 1,965 words ($SD = 1,585.03$; Min. = 622; Max. = 6,827). The thought protocols represented a direct and detailed report of interpretations during the program, and formed the basis to assess interpretive differentiation. For the current test analysis we report on in this article, we used the protocols of all nineteen participants regarding one single news item on political solutions to the 'BSE' or mad cow disease problems in Europe (length: 2:54 min.). Furthermore, we used the two participants' protocols regarding the entire bulletin, which we expected, on face value, to have either relatively complex or simple interpretations.

Data analysis: Three phases

Measuring interpretive differentiation requires distilling the number of different elements and the range of these elements from the thought protocols. In the following we elaborate on how we operationalized these concepts, and how we constructed a categorization system and coding strategy. The data analysis phase consisted of three phases: construction of basic sentences, coding of elements, and assessment of differentiation scores.

1. *Constructing basic sentences.* Since the way in which participants formulate their thoughts can sometimes be quite diffuse, we broke up each protocol into 'basic sentences'. Each basic sentence represented only one statement loosely based on the structure 'object $x \rightarrow relation \rightarrow$ subject y ' (cf. Kleinnijenhuis, Oegema, De Ridder, and Ruigrok, 1998; Osgood, Sporta, and Nunnally, 1956; Van Cuilenburg, Kleinnijenhuis, and De Ridder, 1988). From these basic sentences, the number and range of elements were distilled². Working with basic sentences proved superior to working directly on the protocols, facilitating coding as it severely reduced the number of elements that remained hidden in diffuse formulations.

2. *Coding of elements.* In order to establish what parts of the basic sentences could be regarded as elements, we used a list of nine categories of general elements composed by Spradley to aid researchers in recognizing interpretive elements (Spradley, 1979, 1980; cf. Schaap et al., 2005). This list represents all types of elements of which an interpretation of any social phenomenon can consist. We regarded these general types of elements as representing the range of elements. The kinds of elements people can use in their interpretation of television news are kinds of actors, goals, feelings, attributions of actors, goals and feelings, causes of actors, etc. The more of these different kinds people use, the larger the range of their interpretation. The specific instances of general types that we encountered in the protocols (e. g., George W. Bush as a specific instance of a kind of actor) were regarded as specific elements. That is, every single specific element a person uses contributes to the number aspect of their interpretation. Each specific element people use in interpreting should fit one of these general types:

Category 1. Inclusion elements

Types of elements: *kinds of* actors, goals and feelings, acts, activities, and events, space, time, and objects.

Category 2. Attribution elements

Types of elements: *attributes of* actors, goals and feelings, acts, activities, and events, space, time, and objects.

Category 3. Rationale elements

Types of elements: *reasons for* actors, goals and feelings, acts, activities, and events, space, time, and objects.

Category 4. Function elements

Types of elements: *functions of* actors, goals and feelings, acts, activities, and events, space, time, and objects.

Category 5. Sequence elements

Types of elements: *steps or phases in* actors, goals and feelings, acts, activities, and events, space, time, and objects.

Category 6. Cause-effect elements

Types of elements: *causes of* actors, goals and feelings, acts, activities, and events, space, time, and objects.

Category 7. Location-for-action elements

Types of elements: *places to do/for* actors, goals and feelings, acts, activities, and events, space, time, and objects.

Category 8. Means-end elements

Types of elements: *ways to do/be* actors, goals and feelings, acts, activities, and events, space, time, and objects.

Category 9. Spatial elements

Types of elements: *parts of* actors, goals and feelings, acts, activities, and events, space, time, and objects.

To develop a complete and workable coding scheme we started by reading our protocols using the above list as a 'prototype coding scheme'. Our goal was to create a coding scheme that was tailored to the elements that were used in the interpretation of this specific news bulletin, which may be different from the elements in the original list. The general types of elements were treated as 'sensitizing concepts'; i. e., each time we classified a term into a category of elements, previously classified similar elements were revisited to analyze their fit or difference. Thus, the coding scheme was under constant scrutiny and subject to change (cf. Glaser and Strauss, 1967).

We classified elements by systematically asking questions related to that specific category, for example, is this term in this protocol a kind of actor, feeling, or object? Is it an attribute of an act, event, or place? This way, we coded the discrete elements, simultaneously developing

concrete descriptions of elements for use as coding instructions (describing: 'what is an actor?', 'what is an object?', etc.). After several rounds of carefully reading and re-reading the basic sentences in this manner, we found that participants in the interpretation of this news bulletin used elements in only six of the nine categories provided by Spradley³. Furthermore, we combined categories 3 and 4, 'rationale' and 'function', for practical reasons. This resulted in a coding scheme containing five categories: inclusion, attribution, cause-effects, rationale and function, and sequence. In these large categories, the smaller types of elements were located (actors, acts, events, objects, etc.), including basic descriptions of each type. In sum, the number of types of elements included in a thought protocol establishes the range of elements used in the interpretation, whereas the number of specific elements in these general types represents the number of elements.

Using the definitive coding scheme, all basic sentences of all participants regarding the BSE news item as well as the basic sentences of two participants related to all news items were re-read and coded. We coded all direct references to, for instance, persons or objects, whenever we were able to assess to whom or what the participant was referring. We did so, even if an actual name was not explicitly mentioned, which was often the case when people simply referred to 'he', 'she', or 'it'⁴. Virtually all statements of our participants could be coded by means of our coding scheme⁵. We established coding reliability using an independent coder trained to use the coding scheme, who practiced coding on 10 protocol segments. This coder coded a random sample of 20% of all protocols. Inter-coder agreement between this coder and the original coder was calculated for exact code agreement. Scott's *pi* for inter-coder agreement was .88 (Scott, 1955).

3. *Assessing differentiation scores.* Table 2 provides a breakdown of the coding process and outcome of the thought protocol of one segment. It also illustrates how scores for number and range of elements were assigned. Each time a coder encountered an individual case (e. g., a specific actor), the code for the corresponding category was assigned, thus establishing a term as one 'element' of a certain type. To establish the 'number' of discrete elements, terms that referred to the same individual case (e. g., the same actor) were counted as only one discrete element⁶. For example, in Table 2 the participant mentions a female politician a number of times ('she'). Thus, a new element was counted whenever a participant mentioned a new, previously not mentioned specific actor, act, object, etc. Range was calculated by adding the number of different types of elements used.

Table 2. *A protocol related to one news segment, basic sentences and codes (Item BSE).***Protocol segment (participant no. 1)***"Well, I thought she had a nice purple shirt.**And what she said about the consumer ... the consumer has absolutely no idea that the minister's policy is erratic.**I think she wants 15 million people to support her, or something, because you won't make it with just a purple shirt."*

Basic sentences	Codes
She ¹ had a nice purple ² shirt ³	1. kinds of actors 2. attributes of objects 3. kinds of objects
What she ¹ said ⁴ about the consumer ⁵	4. kinds of acts 5. kinds of actors
The consumer ⁵ has no idea ⁶ that the minister's ⁷ policy ⁸ is erratic ⁹	6. kinds of feelings 7. kinds of actors 8. kinds of objects 9. attributes of objects
She ¹ wants ¹⁰ 15 mln. people ¹¹ to support her ¹⁰	10. kinds of goals 11. kinds of actors
She wants ¹⁰ 15 mln. people ¹¹ to support her because you won't make it ¹² with just a purple ² shirt ³	12. reasons for goals
Number of elements	12
Range of elements	7

In total, the participant in Table 2 used twelve distinct elements, making the amount of differentiation score – for this segment only – twelve. The range of differentiation score for this segment was seven as the specific elements used were of seven different types.

We obtained some preliminary indications of the validity of our instrument, using educational level as an indicator, since this is conceptually related to cognitive complexity. Previous studies show positive empirical relations between education and cognitive complexity – the differentiation and integration of cognitive structures (Luskin, 1990) – and between education and the level of news processing, represented by recall and comprehension levels and a more complex verbal reproduction (Findahl and Höijer, 1981, 1985; Giegler and Ruhrmann, 1990; Renckstorf and Rohland, 1980; Robinson and Davis, 1990; Robinson and Levy, 1986)⁷. Therefore, educational level can be used as an indirect indicator of the validity of our measurement of differentiation (i. e., construct validity: Cronbach and Meehl, 1955; Zeller and Carmines, 1980). Accordingly, a positive relation between educational level and number of elements and of categories of elements should indicate a valid measurement of differentiation. We compared the number and range of elements used by participants with both a high and a lower education in the inter-

pretation of one news item. Higher educated participants scored, on average, 52.9 on number of elements ($SD = 21.1$) and 32.7 on range ($SD = 13.3$), whereas the lower educated scored on average 30.9 on number ($SD = 16.5$) and 18.9 on range ($SD = 9.1$). A Mann-Whitney test showed that the differences for the number of elements were significant at $\alpha = .05$ (2-tailed), but not for range, although the distribution showed the same trend. This provides us with some provisional validation of our instrument.

Results

As said, we conducted a small pilot study to test the practicality of the study as well as the use of the data it generated. The results are presented here for illustrative purposes. To demonstrate the usefulness of the method and its data, we will assess whether interpretation differences can be found between different viewers. Different viewers have different social, situational, and psychological characteristics represented in different knowledge structures. Therefore, we can assume that they interpret identical television news content differently. Since our research group consisted of participants with at least some variance in three different characteristics (sex, age, and educational level), we would expect differences between participants in the number and range of elements they used in interpreting a news item.

Number

Viewers have, and are willing and able to report, thoughts while watching a news program. As seen from the lengths of the protocols in regard to the number of words (see above), they reproduce quite a large number of thoughts during the whole news program. On average participants incorporated 39 unique elements in their interpretations regarding a single three minute news item on the mad cow disease. There are considerable differences between participants in the number of elements used, with a maximum of 80 and a minimum of 10 elements used (Table 3). In other words, some viewers' interpretations are far more specific than others.

Range

Not only do viewers use a large number elements in interpreting the news, the elements are also often fairly heterogeneous. This indicates that many viewers do not merely think about, for instance, the people, places, or events in a news item, but that they include a wide variety of

Table 3. *Number and range of elements used while watching per viewer (Item BSE).*

Participants	Number (Elements)	Range (Types)
1	80	14
2	66	13
3	46	14
4	32	12
5	38	13
6	42	15
7	34	14
8	23	8
9	10	6
10	78	18
11	42	17
12	39	15
13	25	8
14	14	9
15	63	13
16	23	9
17	49	16
18	14	9
19	23	10
<i>N</i> = 19	<i>N</i> 741 <i>M</i> 39.00 <i>SD</i> 20.82	<i>N</i> 233 <i>M</i> 12.26 <i>SD</i> 3.39

elements in their reconstructions of a news item. Again, there were noticeable differences between participants, although the differences do not seem as extreme as the differences in the number of elements (Min. = 6, Max. = 18). These less extreme differences are partly explained by the fact that the number of elements people may use is unlimited, whereas the number of types of elements is restricted by the number of types in our classification scheme.

Differences between viewers

Using somewhat crude criteria, we divided the participants into a typology of interpretive differentiation, establishing three differentiation profiles for this specific news item (Table 4)⁸. Fisher's exact test for this distribution was significant at $\alpha = .001$. The two foremost groups are the groups with either a high number and range or a low number and range. So, most participants who used many elements did so in many different categories (i.e., their interpretations are both specific and heterogeneous), whereas most participants who used a small number of elements did so in a low variety of categories (unspecific and homog-

enous). The interpretations of these two groups can be called 'differentiated' and 'undifferentiated' respectively. However, a third, although small, differentiation type was also found: interpretations containing relatively few elements (unspecific) in a relatively wide range of categories (heterogeneous). The opposite, interpretations containing many elements of low variety (specific and homogenous), at least in this group did not occur, although it is certainly not inconceivable that such interpretations do exist.

Table 4. *A typology of interpretive differentiation: Number of participants per subgroup (Item BSE).*

Range	Number		
	Low	High	
Low	8	–	8
High	2	9	11
	10	9	<i>N</i> = 19

Differences within viewers

Above, we discussed the indications of differences between participants regarding their interpretation of the same news item. We can hypothesize that these differences can be attributed to personal characteristics such as educational level. However, we can also expect that interpretations of various news items vary within the same viewer. The interpretive differentiation of a participant should presumably be somewhat consistently high or low as it is influenced by structural social and biographical characteristics such as education. Yet, at the same time, interpretive differentiation may be variable. Different news items often concern different knowledge domains, and viewers' knowledge and interests regarding these domains may vary, which in turn may influence their interpretations. To illustrate this, we analyzed interpretive differentiation for two participants. For easy comparison, we selected one participant who seemed, on face value, to have a highly differentiated interpretation of most of the news, and another who seemed to show less overall interpretive differentiation (Table 5). In the following, these two participants, corresponding to participants 1 and 14 in Table 3, are called participant A and B, respectively.

Because we were interested in variations in differentiation between items, we compared five out of seven items with about the same length,

Table 5. *Amount and range of elements used by two viewers per news item.*

News Item and Length	No. of 'breaks'	Participant A		Participant B	
		Number	Range	Number	Range
2. 2:54	10	80	14	12	9
3. 2:36	8	77	18	26	12
4. 2:30	8	112	18	21	8
5. 2:42	10	103	19	22	6
6. 2:24	9	64	13	8	5
<i>N</i>		441	82	89	40
<i>M</i>		88.20	16.40	17.80	8.00
<i>SD</i>		19.41	2.70	7.50	2.74

all ranging from 2:24 to 2:54 minutes, as these items provide the participants with more or less equal amounts of time to think, as well as roughly the same number of content to think about.

As expected, interpretive differentiation of the same viewer regarding different news items varies (Table 5). Maximum scores may be up to 1.5 or 3 times higher than minimum scores within the same viewer (cf. number participant A; number and range participant B). Although this is evidence of some variance within participants in both the number and range between items, participant A consistently showed a higher interpretive differentiation than participant B. On average the number of elements in the interpretation of participant A was almost five times higher than in participant B's interpretation. The range of A's interpretation was about two times higher than B's interpretation. The extent of the difference between these two participants is illustrated by the fact that the maximum number of elements used by participant B is less than half the minimum number used by participant A, and the maximum range score of B is still lower than A's minimum range. Additionally, neither the length of the items nor the number of 'thought-listing breaks' seems to be directly related to the level of differentiation, in that a longer item and more breaks do not necessarily lead to a higher level of differentiation. In sum, differentiation does vary with the news items, yet it also seems relatively consistent within participants, which is consistent with our expectations.

Conclusions and discussion

The aim of this contribution was to outline a method and procedure for the study of television news interpretation and to assess the usefulness of the data it generates for exploring the interpretation of television news

by its viewers. It appears that viewers have a large number of thoughts during the viewing of the news. These thoughts are of varying degrees of specificity and heterogeneity, presumably related to social and personal characteristics among which educational level. Finally, viewers seem to interpret each news item with a different level of differentiation.

These results indicate that the procedure is useful for studying interpretive differentiation. Differentiation scores allow us to distinguish between television news viewers with highly differentiated interpretations and viewers with less differentiated interpretations, as well as between the interpretations of several news items within individual viewers. As interpretive differentiation is presumably strongly related to integration – the second characteristic of interpretive complexity (cf. Schaap et al., 2005) – we expect the instrument to be fit to measure differences in interpretive complexity as a whole. The results corroborate with earlier research that has shown that different viewers of expository programs have different thought profiles, and that different parts of a program relate to different types of thoughts (Findahl and Höijer, 1981, 1985; Höijer, 1989; Renckstorf and Rohland, 1980).

A problematic issue in this study is of course the representativeness of the analysis of thoughts. We have provided a modest test of the validity of our instrument by demonstrating a relation between two theoretical concepts (i. e., educational level and complexity). However, this validation is merely preliminary and should be further explored. In future research the classes of thoughts we have assessed should be tested. Classifications can be supplemented and validated for instance by conducting focused interviews with participants, in which they are asked to judge the researcher's classification, or to provide their own classifications (Spradley, 1979, 1980; Van der Veer, Ommundsen, Hak, and Larsen, 2003). In addition, results of qualitative studies can be used in quantitative research in which classifications can be validated on a larger scale.

How can this method contribute to the study of what people do with the news? When compared to 'classical' studies of television news processing that have used recall and comprehension measures, this is an attempt to do more justice to the complexity that is involved in watching, understanding, and giving meaning to the news. It allows us to study a wider range of the reconstructions viewers make of the news than before, as oftentimes the analysis of recall and understanding of news has been limited to certain specific, expected, and deemed as important, news 'facts'.

Moreover, studying the structure of viewers' interpretations does not require a pre-fixed definition of what constitutes the processing of news content. Therefore, it can give us not only a more detailed but also a less 'biased' view of what people 'do' with the news (cf. Hendriks Vet-

tehen, Schaap, and Schlösser, 2004; Massey, 1995; Renckstorf and Wester, 2001; Schaap, 2004). To illustrate how focusing on the structure of interpretation (which is what we did here) differs from focusing on its content, consider the following example. The statement ‘a fork is used to eat soup’ may, when judged on its content not necessarily be a ‘true’ statement; in reality a fork is hardly ever used to eat soup. In most classical news studies, this would have meant that the statement was classified as ‘false’ or ‘misunderstood’. However, the structure of this statement indicates that the viewer has incorporated in his or her interpretation the function of forks; ‘functions of forks’ are part of his or her interpretation of a news item. The interpretation is ‘true’ to the viewer, regardless of what reality or a researcher may think of it. Moreover, whereas one viewer considers the function of something, and connects this to the news item that is watched, other viewers may include in their reconstruction of the item only the fork, and not its function, or include altogether different kinds of elements. Measuring such structural characteristics, in other words, can give us insight in interpretational differences between viewers without the necessity of making the ‘true/false’ judgments that have been used in much previous research on recall and comprehension of news (Robinson and Davis, 1990). This method takes into account the subjective nature of interpreting the news. Moreover, it allows us to do so fairly systematically.

In addition to the study of viewers’ interpretation, one could apply the concept and operationalization of ‘complexity’ to the analysis of news contents, as they themselves are reconstructions too. Comparisons between the complexity of the news and that of viewers’ interpretations could for instance test hypotheses about ‘bottom-up’ and ‘top-down’ processing of news. Furthermore, relations between viewer characteristics and the complexity of interpretations could be studied. Thus, further studies with similar methods may provide us with supplementary insight into the relationship between interpretation and the influence of television news.

Acknowledgements

The authors would like to thank Solange Schlösser for help in gathering the data, and Anna Bosman for her careful reading, useful suggestions, and corrections to earlier versions of this article.

Notes

1. The research group consisted of ten women and nine men. Age ranged from 20 to 64 years ($M = 38$ years). Twelve participants had completed a low educational level (i.e., any degree up to and including higher vocational education) and seven participants had attained a high education (bachelor’s degree or higher).

2. Not all the words people used when verbalizing their thoughts were equally relevant to measuring interpretive differentiation. The aim was not, for instance, to make a linguistic analysis of the protocols. This meant that for instance frequently used words such as 'a', and 'the' were irrelevant, while other words signified elements used in the interpretation.
3. The categories that were not used were 'location-for-action', 'means-end', and 'spatial'. It is conceivable why these categories have been found in ethnographic research and not in protocols of spontaneous thoughts during a television news item. Ethnographers aim to describe every aspect of a culture, including all its cultural acts and objects. Their strategy is to ask members of a culture to describe all the aspects of interest, and keep asking focused questions until they have all the information they need. One can understand that television news viewers do not spontaneously think about different ways to use a location, or the different parts of something at the moment they are watching the news, unless they are pressed for it by an interviewer.
4. A certain level of context sensitivity was required for this procedure. Sometimes we could classify a statement only if statements in its direct vicinity or even in the news content itself were considered. For instance, in the event a reference was made to a person without directly mentioning a name (e.g., 'she' in Table 2) the coder could deduce from subsequent statements and the segment in the news to which this statement related that 'she' was the politician making a speech in the segment. In almost all cases, the coder was able to deduce without much doubt to whom or what participants referred. In cases where this was impossible, we classified the element in a 'missing category'.
5. The only exceptions were statements that are perhaps typical of a news-watching situation, as opposed to interviews. These include short cries and exclamations such as 'well, well', 'how about that', and 'gee-whiz' which may occur without any further context. Without even the smallest context, it becomes impossible to classify them. When the coder encountered such statements, and the context of the rest of the statements did not provide the coder with a clear idea of how to categorize the statement, the coder placed them in a 'not codable' category.
6. This is another instance in which sensitivity to the context of the entire statement and news item is important to determine whether a statement refers to the same actor or not.
7. Both the theoretical and the empirical relation between education and cognitive complexity and news processing is indirect, via IQ and interest (Graber, 1984; Luskin, 1990) as well as cognitive skills, level of processing (Woodall et al., 1983), and the level of specific and general knowledge (Findahl and Höjjer, 1985; Giegler and Ruhrmann, 1990). Therefore, correlation with educational level is only a very indirect indicator of construct validity.
8. The mean scores were taken as a criterion for classifying participants into either the high or low groups; for number we placed participants with the score 39 in the 'high' category.

References

- Cronbach, L. J. and Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Cuilenburg, J. J. van, Kleinnijenhuis, J., and Ridder, J. A. (1988). *Tekst en betoog* [Text and argument]. Muiderberg: Coutinho.

- Findahl, O. (2001). News in our minds. In K. Renckstorf, D. McQuail, and N. Jankowski (Eds.), *Television news research: Recent European approaches and findings* (pp. 111–128). Berlin: Quintessence.
- Findahl, O. and Höijer, B. (1981). Studies from the perspective of human comprehension. In C. Wilhoit and H. de Bock (Eds.), *Mass Communication Review Yearbook* (pp. 393–403). Beverly Hills, CA: Sage.
- Findahl, O. and Höijer, B. (1985). Some characteristics of news memory and comprehension. *Journal of Broadcasting and Electronic Media*, 29, 379–396.
- Giegler, H. and Ruhrmann, G. (1990). Remembering the news: A LISREL model. *European Journal of Communication*, 5, 463–488.
- Glaser, B. G. and Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago, IL: Aldine.
- Graber, D. K. (1984). *Processing the news: How people tame the information tide*. New York: Longman.
- Hendriks Vettehen, P., Schaap, G., and Schlösser, S. (2004). What men and women think while watching the news: An exploration. *Communications: The European Journal of Communication Research*, 29, 235–252.
- Höijer, B. (1989). Television-evoked thoughts and their relation to comprehension. *Communication Research*, 16, 179–203.
- Höijer, B. (1990). Studying viewers' reception of television programmes: Theoretical and methodological considerations. *European Journal of Communication*, 5, 29–56.
- Kleinnijenhuis, J., Oegema, D., De Ridder, J. A., and Ruigrok, P. C. (1998). *Paarse polarisatie: De slag om de kiezer in de media* [Purple polarization: The battle for the voter in the media]. Alphen a/d Rijn: Samsom.
- Luskin, R. C. (1990). Explaining political sophistication. *Political Behavior*, 12, 331–361.
- Massey, K. B. (1995). Analyzing the uses and gratifications concept of audience activity with a qualitative approach: Media encounters during the 1989 Loma Prieta earthquake disaster. *Journal of Broadcasting and Electronic Media*, 39, 328–349.
- Osgood, C. E., Sparta, S., and Nunnally, J. C. (1956). Evaluative assertion theory. *Litera*, 3, 47–102.
- Renckstorf, K. and Rohland, L. (1980). *Nachrichtensendungen im Fernsehen* [News broadcasts on television]. Berlin: Spieß.
- Renckstorf, K. and Wester, F. (2001). An action theoretical frame of reference for the study of television news use. In K. Renckstorf, D. McQuail, and N. Jankowski (Eds.), *Television news research: Recent European approaches and findings* (pp. 91–110). Berlin: Quintessence.
- Robinson, J. P. and Davis, D. K. (1990). Television news and the informed public: An information-processing approach. *Journal of Communication*, 40, 109–116.
- Robinson, J. P. and Levy, M. R. (1986). *The main source: Learning from television news*. Beverly Hills, CA: Sage.
- Robinson, J. P. and Levy, M. R. (1996). Media and the informed public: A 1990s update. *Journal of Communication*, 46, 129–137.
- Schaap, G. (2004). Using protocol analysis in television news research: Proposal and first tests. In K. Renckstorf, D. McQuail, J. E. Rosenbaum, and G. Schaap (Eds.), *Action theory and communication research: Recent developments in Europe* (pp. 115–140). Berlin: Mouton-De Gruyter.
- Schaap, G., Renckstorf, K., and Wester, F. (2001). Three decades of television news research: An inventory of issues and problems. In K. Renckstorf, D. McQuail, and N. Jankowski (Eds.), *Television news research: Recent European approaches and findings* (pp.47–90). Berlin: Quintessence.

- Schaap, G., Renckstorf, K., and Wester, F. (2005). Conceptualizing television news interpretation by its viewers: The concept of interpretive complexity. *Communications: The European Journal of Communication Research*, 30, 269–291.
- Scott, W. A. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly*, 19, 503–515.
- Spradley, J. P. and McCurdy, D. W. (1972). *The cultural experience: Ethnography in complex society*. Chicago, IL: Science Research Associates.
- Spradley, J. P. (Ed.). (1972). *Culture and cognition: Rules, maps, and plans*. San Francisco, CA: Chandler.
- Spradley, J. P. (1979). *The ethnographic interview*. New York: Holt, Rinehart and Winston.
- Spradley, J. P. (1980). *Participant observation*. New York: Holt, Rinehart and Winston.
- Someren, M. W. van, Barnard, Y. F., and Sandberg, J. A. C. (1994). *The think aloud method: A practical guide to modeling cognitive processes*. London: Academic Press.
- Veer, K. van der, Ommundsen, R., Hak, T., and Larsen, K. S. (2003). Meaning shift of items in different language versions: A cross-national validation study of the Illegal Aliens Scale. *Quality and Quantity*, 37, 193–206.
- Woodall, W. G., Davis, D. K., and Sahin, H. (1983). From the boob tube to the black box: Television news comprehension from an information processing perspective. *Journal of Broadcasting*, 27, 1–23.
- Zeller, R. A. and Carmines, E. G. (1980). *Measurement in the social sciences*. Cambridge: Cambridge University Press.