Toward an empirical verification of the General Theory of Verbal Humor

Ruch, Willibald; Attardo, Salvatore; Raskin, Victor

Abstract: The present study derives hypotheses from the General Theory of Verbal Humor (GTVH) and tests them on a sample of 534 subjects. Subjects are presented with three sets of jokes, each consisting of an anchor joke and comparison jokes in which variations in one and only one of the six Knowledge Resources (KR), script Opposition (SO), logical mechanism (LM), Situation (SI) target (TA), narrative strategy (NS), andlanguage (LA) occurred. Subjects rated the degree of similarity between the anchor joke and the six comparison jokes. The results support the hypothesis that the extent to which the similarity judgment is affected depends on the type of the KR manipulated. Also, there generally is a decreasing trend in similarity between the KRs LA and SO. Whereas there was a significant difference between all consecutive KRs, äs predicted by the hierarchy postulated by the GTVH, SI and LM were not in the right order. Possible explanations for thisfact are discussed. This article presents and discusses a study which empirically Supports some of the Claims of the General Theory of Verbal Humor (GTVH). After introducing the theory, the article will present the hypotheses derived from the theory that were tested and finally the results of the investigation.

DOI: https://doi.org/10.1515/humr.1993.6.2.123

Toward an empirical verification of the General Theory of Verbal Humor

WILLIBALD RUCH, SALVATORE ATTARDO, and VICTOR RASKIN

Abstract

The present study derives hypotheses from the General Theory of Verbal Humor (GTVH) and tests them on a sample of 534 subjects. Subjects are presented with three sets of jokes, each consisting of an anchor joke and comparison jokes in which variations in one and only one of the six Knowledge Resources (KR), script opposition (SO), logical mechanism (LM), situation (SI) target (TA), narrative strategy (NS), and language (LA) occurred. Subjects rated the degree of similarity between the anchor joke and the six comparison jokes. The results support the hypothesis that the extent to which the similarity judgment is affected depends on the type of the KR manipulated. Also, there generally is a decreasing trend in similarity between the KRs LA and SO. Whereas there was a significant difference between all consecutive KRs, as predicted by the hierarchy postulated by the GTVH, SI and LM were not in the right order. Possible explanations for this fact are discussed.

This article presents and discusses a study which empirically supports some of the claims of the General Theory of Verbal Humor (GTVH). After introducing the theory, the article will present the hypotheses derived from the theory that were tested and finally the results of the investigation.

The SSTH and GTVH

Raskin (1985) presented in detail the first formal semantic theory of jokes, from which the GTVH derives. Given its reliance on the concept

Humor 6–2 (1993), 123–136. 0933–1719/93/0006–0123 $2.00

of "script" (a structured chunk of information about lexemes and/or parts of the world), Raskin's theory became known as the Semantic Script Theory of Humor (SSTH). The SSTH can be summarized as two necessary and sufficient conditions for a text to be funny:

(1) a. Each joke must contain two overlapping scripts (that is, the joke must be interpretable, fully or in part, according to two different scripts);
b. The two scripts must be opposed (that is, they must be the negation of each other, if only for the purpose of a given text), according to a list of basic oppositions, such as real/unreal, possible/impossible, etc.

Attardo and Raskin (1991) revised the SSTH, and the result of this revision was the General Theory of Verbal Humor (GTVH). The revision of the SSTH consisted mostly of broadening its scope by the introduction, besides scripts, of five other Knowledge Resources (KR) that must be tapped into when generating a joke. The KRs are script opposition (SO), logical mechanism (LM), situation (SI), target (TA), narrative strategy (NS), and language (LA). The GTVH also focused on joke similarity and dedicated a great deal of effort to establishing the concept formally (for details, see Attardo and Raskin 1991).

We will briefly describe the six KRs and then discuss the concept of joke similarity.

Language (LA)

The LA KR is the actual verbalization of the joke, resulting in its text. It includes all the linguistic components of the text at all levels.

Narrative strategy (NS)

The NS KR accounts for the fact that any joke has to be cast in some form of narrative organization, that is either as a simple (framed) narrative, as a dialogue (question and answer), as a (pseudo-)riddle, as an aside in a conversation, etc.

Target (TA)

The target KR selects the butt of the joke. Jokes that are not aggressive (that is, that do not ridicule someone or something) have an empty value for the TA.

Situation (SI)

Any joke must introduce some event or situation such as changing a light bulb, crossing the road, playing golf, etc. The situation of a joke can be thought of as the "props" of the joke: the objects, participants, instruments, activities, etc.

Logical mechanism (LM)

The logical mechanism accounts for the way in which the two senses (scripts) in the joke are brought together. LMs can range from straightforward juxtapositions, as in the tee-shirt slogan reading

(2) Gobi Desert Canoe Club

to more complex errors in reasoning, such as false analogies, garden-path phenomena, as in

(3) Madonna does not have one, the Pope has one but doesn't use it, Bush has a short one, and Gorbachev has a long one. What is it? Answer: a last name.

or figure-ground reversals, as in

(4) How many Poles does it take to screw in a light bulb? Five. One to hold the light bulb and four to turn the table he's standing on. (light bulb: figure; body: ground)

Script opposition (SO)

This KR deals with the script opposition/overlap requirement presented in the SSTH. It should be noted that the SO is the most abstract (perhaps

sharing this degree of abstractness with the LM) of all the KRs, which accounts for the fact that the SSTH could collapse all six KRs onto this one (basically ignoring the other five, with some exceptions, such as TA and LA).

The joke, according to the GTVH

From the point of view of the GTVH, each joke can be viewed as a six-tuple, specifying the instantiation of each KR as a parameter:

(5) Joke: {LA, SI, NS, TA, SO, LM}

The GTVH presents itself as a mechanism capable of generating an infinite number of jokes by combining the various values that each parameter can take. It should be noted that these values are not binary. The values for the LM and the SO seem to be limited in number (see, respectively, Attardo [1988: 357], and Raskin [1985: 127]), while the SI and LA are much more numerous. Using this powerful mechanism, a taxonomy of jokes has been elaborated upon (Raskin and Attardo 1991) which can assign a unique descriptor to any joke.

A highly technical aspect of the GTVH is the issue of the ordering of the KRs. A complete discussion would be out of place in this context. It is enough to say that various considerations of interdependence and/or independence among the KRs have allowed the postulation of the hierarchical organization in Figure 1.

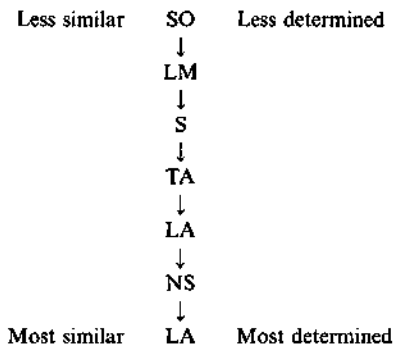

Figure 1. *Hierarchical organization of the KRs*

The basic principle on which the hierarchy is based is that a KR is likely to determine or be determined by another KR. "Determination" is to be intended as limiting or reducing the options available for the instantiation/actualization of the parameter. Consider for example the SO KR. If we select the "dumb/smart" SO, this will determine our choice of TA (Poles, Dan Quayle, and other groups or individuals for which the mythical scripts of dumbness are available). On the other hand, a joke about Poles could be about stupidity but also about other traits (ignorance, etc. that may happen to be stereotypically available for that group). Thus, since the choice of the SO determines a choice in the TA, whereas the opposite is not the case, we will say that the SO is independent from the TA and that the TA depends on SO. Accordingly, the SO will be higher in the hierarchy and the TA lower (Figure 1). This procedure was applied in painstaking detail in Attardo and Raskin (1991), and the outcome was the hierarchy in Figure 1. It should be noted that the arrangement of some KRs was presented as very tentative, most notably between the SO and LM (see below).

On the basis of the degree of determination among KRs (Figure 1, right column), a first extension of the GTVH was applied and it was assumed that determination correlated directly with joke similarity (Figure 1, left column). Thus, it was postulated that a scale of similarity would match entirely the postulated scale of cross-determination. The basis for this assumption is the intuitive fact that the less determined a KR is, the more it allows the introduction of diverging elements that will cause a perception of dissimilarity. In other words, it was assumed that if two jokes differed in only one KR, the difference would be greater if the KR were higher in the hierarchy (less determined and thus more open to free variation) and smaller if the KR were lower in the hierarchy.

The present study

Based on the GTVH, two hypotheses can be formulated as to the behavior of speakers faced with a humorous text:[1] first, the subjects will perceive some jokes as more similar and other jokes as less similar to one another; second, if the GTVH is correct, subjects will perceive a linear[2] increase of similarity between pairs of jokes selected along the KR hierarchy. This second hypothesis includes the following five subhypotheses:

1. jokes differing in LA are more similar than jokes differing in NS;
2. jokes differing in NS are more similar than jokes differing in TA;
3. jokes differing in TA are more similar than jokes differing in SI;
4. jokes differing in SI are more similar than jokes differing in LM; and
5. jokes differing in LM are more similar than jokes differing in SO

as well as the assumption of transitivity.

Thus, the GTVH predicts that given a joke A and a joke B differing, say, in a "low" parameter like LA and a pair of jokes C/D differing in a "high" parameter like SO, the first pair will be perceived as more similar and the second as less similar.

Taking one joke as basic (the "anchor joke"), we generated a set of six jokes, all differing in one KR. (See the three sets of jokes generated with this procedure in Appendix A.) The first joke of each set is the invariant (the anchor joke); the other six jokes all differ from the anchor joke by one and only one KR. The members of this set can then be ordered on the basis of the hierarchy, and this order is presumed to reflect their greater or lesser similarity to the anchor joke.

The GTVH can be interpreted as predicting that perception of similarity is a linear function. Since the KRs are ordered linearly, with the KR causing the least difference in a pair of jokes at the bottom and the KR causing the maximum difference at the top of the hierarchy, we can predict that the perception of similarity will also be represented by a line. Graphically, if we represent the overall prediction of the GTVH as a graph having on the x-axis the hierarchy, with the KR hierarchy arranged on the axis from left (LA) to right (SO), and on the y-axis the perceived degree of similarity (lowest being least similar and higher most similar), the GTVH predicts that the subjects' ranking should be represented by a line having its origin in the top left of the graph and ending in the low right of the graph (see Figure 2).

The general prediction can also be interpreted slightly differently, as the general prediction that any pair of jokes differing in a KR higher on the hierarchy will be perceived as less similar than a pair differing in a KR lower in the hierarchy. This leads us to predict that given any pair of jokes differing by two different KRs (for instance, LA and NS) from the same anchor joke, the joke differing from the KR lower in the hierarchy will be higher in our graph (that is, will be more similar).

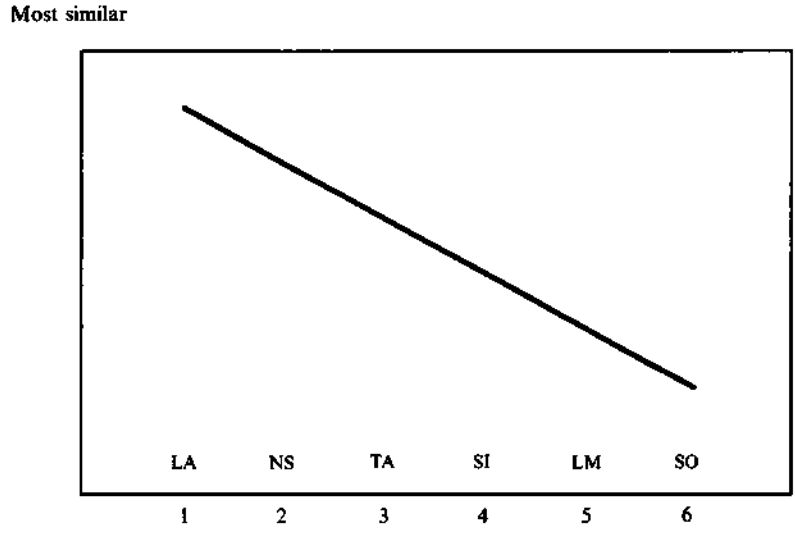

Figure 2. *Projected previsions of the GTVH*

Method

Subjects

To test the GTVH's predictions, we tested the perception of similarity among jokes in two groups of approximately 280 students, which yielded a total population of 534 valid responses, using a battery of similarity perception judgments. The subjects were all undergraduate students enrolled in an introductory psychology class, their participation in the test was voluntary, and they received some scholastic credit for their participation but no monetary or other compensation.

Materials

The materials used in the study consisted of three sets of seven jokes based on six versions of a joke created by manipulating the KRs one at a time from a basic joke. The three sets of jokes were built starting with "blonde" jokes, "light bulb" jokes, and "chicken" jokes on the assumption that the subjects would be familiar, if not with the jokes themselves,

at least with the type of joke in general and thus could focus on the differences between them. The three sets of jokes are presented in Appendix A. The three sets of jokes were numbered and presented in different orders, but the order of presentation did not vary for each set.

Procedure

The test was administered in one single session lasting roughly two hours (the entire test had other questions dealing with Dan Holt's[3] research on humor and giftedness). The subjects were presented with a five-point scale (undecided, somewhat similar, similar, very similar, extremely similar) and were asked to evaluate, according to the scale, nine pairs of jokes. The nine pairs of jokes were presented on the same page by numbers (for instance, jokes 5 and 3), and the subjects marked their answer on a scoring sheet. The jokes were presented in different orders within each set but in the same order for all subjects. The nine pairs included the six relevant comparison pairs plus three filler pairs that were not used for our analysis but that were designed to prevent the subjects from realizing that all six permutations were being tested. The scoring sheets were tabulated by an optical scanner.

The methodological limitations of this investigation are evident, and the possibility of an order effect has not been ruled out by the fixed presentation of the stimuli. However, the three sets of jokes had different orderings, as well as the joke pairs, and thus any significant effect would have been detected (for instance because one different pair would have stood out in each set). Regardless, this investigation was intended as a pilot study to test the overall validity of the GTVH. Further studies, with a complete set of pairs, randomized presentation of the stimuli, and control jokes, are in the making.

Results

The judgments were averaged across the three set of jokes (blonde, chicken, light bulb). A repeated measurement ANOVA confirmed the hypothesis that the subjects perceive the jokes as differently similar, $F(5,2670) = 620.54$, $P < 0.0001$. In other words, naive subjects are able to discriminate among the joke pairs with respect to their similarity. A trend

analysis was performed to test the second hypothesis. The repeated measurement factor was tested only for a linear trend, and this effect turned out to be highly significant, $F(1,2670)=2595.586$, $P<0.0001$. This contrast accounts for 83.66% of the explained variance and thus provides support for the hypothesis of a linear decrease in similarity between the KRs LA and SO. The actual means are given in Figure 3 along with the means for the three sets of jokes.

Figure 3 shows the expected downward trend for the six comparisons. T-tests between consecutive KRs were performed and appeared to be highly significant (all $Ps<0.001$). However, SI and LM are not in the right order; the variation in the situation is perceived as creating more dissimilarity than a variation in the logical mechanism. The mean for SI is even lower than the one for SO; this difference is not, however, significant (unadjusted $P=0.075$). SI not considered, the decrease in similarity is almost perfectly linear; if SI is excluded from the trend analysis, the effect accounts for 96.63% of the explained variance.

Figure 3 shows that the results for the blonde and light bulb jokes follow very much the pattern described above for the average score. There is a more marked decrease between LM and SO, and there is even a decrease between SI and SO as expected by the second hypothesis.

For the chicken joke, however, there are more deviations from the

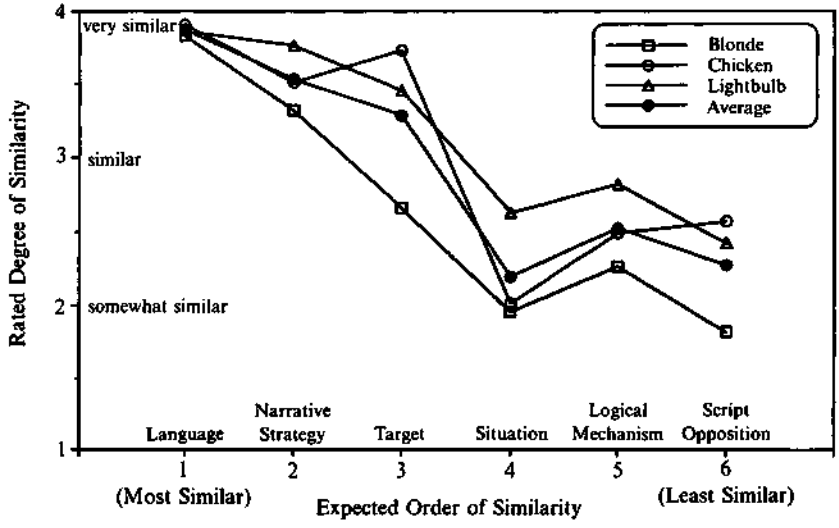

Figure 3. *Rated degree of similarity among the joke pairs*

expectations. There is not only the reversal between SI and LM (as also observed for the blonde and light bulb jokes), but also TA is higher than NS, and SO is higher than LM (although only slightly so). The chicken joke yields the most violations of the expectations, although the general downward trend is still observable (accounting for 61.99% of the repeated measures variance).

Discussion

The goal of the present article was to derive hypotheses from the General Theory of Verbal Humor (GTVH) and to test them empirically using three sets of jokes. The predictions related to the subjects' perception of similarity among joke pairs differing in one Knowledge Resource (KR).

The first hypothesis received full support; the manipulations undertaken with the jokes led to differences among them which were reliably perceived by the subjects. Thus, even untrained laypersons were sensitive to changes in the KRs underscoring the validity of these concepts. The extent to which the similarity judgment is affected depended on the type of KR manipulated; variations in some KRs make the joke pairs more dissimilar than variations in others.

The second hypothesis received support as far as the overall prediction of a decreasing trend in the perception of similarity between the KRs LA and SO is concerned. This general decrease in similarity can be found in all three sets of jokes as well as in the average scores. With regard to the subhypotheses, most of the predictions were confirmed as well. On the average, variation in NS made the joke less similar to the anchor joke than variations in LA (which led to the least amount of difference from the anchor joke). Also, greater dissimilarity is created by variations in TA rather than NS, in SI rather than TA, and SO rather than LM. Contrary to the expectations, however, variation in LM makes a comparison joke less different from the anchor than a variation in SI.

Thus, the perception of similarity across the KRs is as predicted by the GTVH with the exception of the LM (and the SO and TA in the chicken set). There are a number of possible explanations for this fact. The first possibility is that the subjects are being asked to perform a task that is too complex, sophisticated, or simply fine-grained for them to perform as expected. In this case, either a more sophisticated experimental design is required, or one would have to reach the conclusion that the

SI LM SO
TA
NS
LA

Figure 4. *The "T" model*

difference cannot be mapped using introspective questions. The fact that the first hypothesis was supported by the present study, however, seems to rule out this interpretation. The second possibility is that the ordering of the KRs in the GTVH is incorrect and that the lack of consistency in the differentiation between SI, LM, and SO is a sign that the hierarchy should be redesigned, for example, as shown in Figure 4.

There are theoretical objections to this "T" model, discussed in detail in Attardo and Raskin (1991), that make it an undesirable option. However, the model which would give the best fit of the present results would be of a "Y" shape; in this model only SO and SI are located at the highest level of the hierarchy, with LM below them but above TA. A third possibility is related to the fact that, as Attardo and Raskin (1991) highlighted, the LM is the least explored of all KRs and that there could be several factors at play which might have influenced the subjects' perception. This is further confirmed by the tentative relative ordering of SO and LM in Attardo and Raskin (1991).[4] A fourth hypothesis might be that, in fact, the LM is an artifact of the theory and should be removed altogether. However, this claim is not consistent with the perception, on average, of a difference in similarity between the three lowest KRs and LM. For the time being, and in view of a more complete follow-up study, the question will be left open.

Finally, we have to caution that we considered similarity as a unidimensional construct. It might be that subjects would use more dimensions to portray the similarity between the jokes. A complete comparison between all possible pairs of jokes is needed and application of multidimensional scaling techniques will allow the determination of the number and nature of the dimensions involved. If more than one dimension is necessary, hypotheses relating to similarity/dissimilarity between KRs should be determined in this n-dimensional space.

The present study contains a further untested assumption. It is implicitly assumed that *any* variation in a KR leads to the same amount of dissimilarity from the anchor joke. It is not tested yet that different versions of a variation in, for example, the situation are interchangeable.

At the moment, the argument cannot be ruled out that some targets produce more dissimilarity than others. In the introduction, we pointed out that the values for the different KRs are not binary; especially for SI and LA they are much more numerous. It might well be that the discrepant results found for SI in the present study might be due to the selection of variations in this KR which lie on the very end of the possible dissimilarity. The present study concentrated on the mean deviation from the anchor jokes produced by variations in the single KRs; however, the *range* of possible dissimilarity produceable for a given joke by variations in a given KR remains to be tested. It is also not known whether the results of the present study are generalizable across different types of subjects. Would subjects high and low in verbal ability perceive the similarity of structure among the jokes in the same way? Is the effect of the substitution of the target the same for subjects being or not being a member of this particular group? Finally, whereas we know now that variation in some KRs make the joke more dissimilar from the anchor than variations in others, we do not know to what extent the KRs affect the degree of perceived funniness of the jokes. Thus, further research is requested. Despite the obvious limitations of the present study, the results provide support for the theory by confirming most of its predictions.

University of Düsseldorf
Youngstown State University
Purdue University

Appendix

Series A

What do you call it when a blonde dyes her hair brown? Artificial Intelligence. (Anchor)
What's the result of a blonde dyeing her hair brown? Artificial Intelligence. (LA)
When a blonde dyes her hair brown, it's called Artificial Intelligence. (NS)
What do you call it when a fair-haired sorority girl dyes her hair brown? Artificial Intelligence. (TA)
What do you call it when a blonde "lipsyncs" Einstein on the screen? Artificial Intelligence. (SI)

What do you call it when a blonde dyes her hair brown? Illiteracy: she could not read the label on the bottle. (LM)

What do you call it when a blonde dyes her hair brown? Serial murder: her five boyfriends hang themselves. (SO)

Series B

Why did the chicken cross the road? It wanted to get to the other side. (Anchor)

Do you know the reason why the chicken decided to cross the road? Because it wanted to get to the other side. (LA)

The reason the chicken crossed the road is that it wanted to get to the other side. (NS)

Why did the turtle cross the road? It wanted to get to the other side. (TA)

Why did the chicken eat an octagonal-headed worm? Because it was hungry. (SI)

Why did the chicken cross the road? Nothing ventured, nothing gained. (LM)

Why did the chicken cross the road? He saw a blonde hen on the other side. (SO)

Series C

How many Poles does it take to screw in a light bulb? Five. One to hold the light bulb and four to turn the table he's standing on. (Anchor)

The number of Polacks needed to screw in a light bulb? Five — one holds the bulb and four turn the table. (LA)

It takes five Poles to screw in a light bulb: one to hold the light bulb and four to turn the table he's standing on. (NS)

How many Irishmen does it take to screw in a light bulb? Five. One to hold the light bulb and four to turn the table he's standing on. (TA)

How many Poles does it take to wash a car? Two. One to hold the sponge and one to move the car back and forth. (SI)

How many Poles does it take to screw in a light bulb? Five. One to hold the light bulb and four to look for the right screwdriver. (LM)

How many Poles does it take to screw in a light bulb? Five. One to take his shoes off, get on the table, and screw in the light bulb and four to wave the air deodorants to kill his foot odor. (SO)

Notes

1. Although, it is supposed to account, in principle, for any type of humorous text, reasons of simplicity and experimental design have led us to limit ourselves to jokes.
2. In fact, nothing in the GTHV requires the assumption of linearity in the increase of similarity/dissimilarity perception but rather the monotony of the increase, that is, increase only. The assumption of linearity has been made to simplify matters, and the results seem to confirm this non-trivial fact.
3. Dan Holt's role in administering the test is gratefully acknowledged by the authors.
4. As a matter of fact, the data provide us only with a relative comparison, and all they say is that SI and LM are reversed in order; whether this is due to SI or to LM cannot be decided on the grounds of the present data. Nevertheless, given the theoretical status of the two KR, the interpretation presented in the text seems to be favored.

References

Attardo, Salvatore
1988 Trends in European humor research: towards a text model. *HUMOR: International Journal of Humor Research* 1(4), 349–369.

Attardo, Salvatore, and Victor Raskin
1991 Script theory revis(it)ed: joke similarity and joke representation model. *HUMOR: International Journal of Humor Research* 4(3/4), 293–347.

Raskin, Victor
1985 *Semantic Mechanisms of Humor*. Dordrecht: Reidel.

Raskin, Victor, and Salvatore Attardo
1991 Humor space: a multidimensional feature taxonomy of jokes. Paper presented at the Ninth International Conference on Humour and Laughter. Brock University, Ontario, Canada. June 26–30, 1991.