Journal of Experimental Psychology:
Learning, Memory, and Cognition
2005, Vol. 31, No. 2, 374–377

0278-7393/05/$12.00 DOI: 10.1037/0278-7393.31.2.374

COMMENTARY

Predicting Without Modeling: A Critique of Trabasso and Bartolone (2003)

Stefan L. Frank and Mathieu Koppen
Radboud University Nijmegen

Leo G. M. Noordman
Tilburg University

Wietske Vonk
Max Planck Institute for Psycholinguistics and Radboud University Nijmegen

T. Trabasso and J. Bartolone (2003) used a computational model of narrative text comprehension to account for empirical findings. The authors show that the same predictions are obtained without running the model. This is caused by the model's computational setup, which leaves most of the model's input unchanged.

Trabasso and Bartolone (2003) used the computational model of narrative text comprehension developed by Langston and Trabasso (1999; Langston, Trabasso, & Magliano, 1999). In this model, strengths of causal relations between the text's clauses are indicated by *connection strength values* assigned to each pair of clauses. The Langston and Trabasso model adjusts these assigned values, resulting in output values that are to correspond to a reader's interpretation of the text's causal relations. In this commentary, we show that the model is not instrumental in obtaining the simulation results presented by Trabasso and Bartolone.

A story's causal structure is one of the most important factors affecting comprehension of the story. The causal relations of a statement are known to affect its reading time (Myers, Shinjo, & Duffy, 1987), its recall (Myers et al., 1987; Trabasso & Van den Broek, 1985), ratings of its importance (Trabasso & Sperry, 1985), and the reinstatement of story statements in working memory (Lutz & Radvansky, 1997; Suh & Trabasso, 1993). For an overview, see Trabasso (in press). In applying the Langston and Trabasso (1999; Langston, Trabasso, & Magliano, 1999) model, the modeler determines causal relations between the text's clauses (by means of a counterfactual test). Next, these causal relations are encoded in the connection strength values that form the model's input. When based on a valid causal analysis of the text, these initial strengths will certainly account for empirical data to a considerable extent. For the Langston and Trabasso model to be useful, it needs to be shown that its output, that is, the adjusted connection strengths, predict the data better than the initial strengths do. Below, we first show that, in the simulations run by Trabasso and Bartolone (2003), the model's adjustment of strength values makes no difference to the model's predictive value. Next, we argue that this is a structural problem inherent to the Langston and Trabasso model.

Trabasso and Bartolone's (2003) Results Reexamined

Trabasso and Bartolone (2003) investigated a story used in an experiment by Kahneman and Tversky (1982). This story describes a Mr. Jones, who is killed in a car accident while driving home from his office. In one version (the *Route* version), Mr. Jones leaves his office at the regular time but does not take his regular route. Another version (the *Time* version) states that he leaves at an unusual time but takes his regular route. In either version, Mr. Jones stops at a crossing, although he does not need to, and gets hit by a truck driven by a boy under the influence of drugs.

Kahneman and Tversky (1982) had participants read either the Route or the Time version of the Mr. Jones story and asked them how Mr. Jones's family and friends would finish a sentence starting *If only* . . . . The results showed that the participants who read the Route version predominantly referred to the unusual route taken, whereas those who read the Time version predominantly referred to Mr. Jones's unusual time of departure.

Trabasso and Bartolone (2003) applied the Langston and Trabasso (1999; Langston et al., 1999) model to these two story versions. For both versions, they determined between which events direct causal relations existed. Next, each pair of clauses received an initial connection strength that depended on the length of the causal path between the corresponding events. In this way, each text is represented as a network of nodes that correspond to the clauses and connections between them having values that indicate the strength of the causal relation between the clauses. After running the model on the two text networks, updated strength values predicted the participants' answers in the Kahneman and Tversky (1982) study. For instance, the proportion of participants who answered *If only Mr. Jones had taken his normal route* is

Stefan L. Frank and Mathieu Koppen, Nijmegen Institute for Cognition and Information, Radboud University Nijmegen, Nijmegen, The Netherlands; Leo G. M. Noordman, Discourse Studies, Tilburg University, Tilburg, The Netherlands; Wietske Vonk, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, and Center for Language Studies, Radboud University Nijmegen.

Leo G. M. Noordman is now at the Nijmegen Institute for Cognition and Information, Radboud University Nijmegen.

Correspondence concerning this article should be addressed to Stefan L. Frank, Nijmegen Institute for Cognition and Information, Radboud University Nijmegen, P. O. Box 9104, 6500 HE Nijmegen, The Netherlands. E-mail: S.Frank@nici.ru.nl

Table 1
Proportion of Participants Referring to an Event and Corresponding Average Connection Strengths

Story version and event	Observed proportion	Connection strength: Output	Input
Route			
Route	.51	4.24	4.06
Time	.03	0.58	0.50
Crossing	.22	2.74	2.75
Boy	.20	2.61	2.64
Time			
Route	.13	0.63	0.55
Time	.26	3.33	3.12
Crossing	.31	3.04	3.15
Boy	.29	2.61	2.79
r^2		.82	.81

Note. Proportion of participants referring to an event (Kahneman & Tversky, 1982) and average connection strength of the corresponding node, both after running the Langston and Trabasso (1999; Langston et al., 1999) model (output, Trabasso & Bartolone, 2003) and without running the model (input, computed by us from Figure 1 in Trabasso & Bartolone, 2003).

predicted by the average final strength of the connections between the clause about Mr. Jones not taking his normal route and each clause of the text.

In Table 1, we compare the empirical findings of the Kahneman and Tversky (1982) study with the model's output according to Trabasso and Bartolone (2003). The average output strengths account for 82% of the variance in the observed proportions of participants. However, the table also shows that virtually the same percentage of variance (81%) is explained by the networks' initial strengths, which we computed directly from the causal networks presented in Trabasso and Bartolone's Figure 1 without running the Langston and Trabasso (1999; Langston et al. 1999) model. This shows that updating the strengths does not increase the predictive value of the causal network.

According to Kahneman and Tversky (1982), the effect of story version on participants' answers occurs because it is easier to imagine Mr. Jones not performing the unusual action than it is to mentally undo the usual action. However, Trabasso and Bartolone (2003) note that the two texts showed confounding between normality and explanation of actions. Mr. Jones's unusual actions were more elaborated than his usual actions, and this gives an alternative explanation for the effect found by Kahneman and Tversky. In order to separate the two possible causes of the effect, Trabasso and Bartolone constructed four more versions of the story. In these versions, it was either Mr. Jones's route or Mr. Jones's time that was unusual, and the reason for this was either explained or unexplained. After reading these texts, participants were given four *If only . . .* sentences referring to the route taken, the time of departure, stopping at the crossing, or the boy who caused the fatal accident. They were asked to rank them by the likelihood that Mr. Jones's family and friends would have these sentences on their mind. As shown in Table 2, participants gave a higher ranking (i.e., the rank number was lower) to a sentence when the corresponding unusual action was explained compared with when it remained unexplained, thereby confirming Trabasso and Bartolone's alternative explanation.

Next, the text networks of these four story versions were processed by the model. Table 2 shows that the average final connection strength of the network node corresponding to Mr. Jones's regular route in the Route versions predicted the average rank order of the statement *If only he had taken a different route home*. Likewise, the average final connection strength of the node corresponding to his regular departure time in the Time versions predicted the average rank order of the statement *If only he had left at a different time* (in total, r^2 = .88). However, this prediction is no less accurate if the initial strengths are used instead of the updated strengths (r^2 = .89). Again, updating the connection strengths does not improve the networks' predictions.

As is clear from Tables 1 and 2, there is only a small difference between input and output connection strengths. Averaged over the six networks used by Trabasso and Bartolone (2003), we found that 89.5% of variance in final connection strengths was accounted for directly by the initial strengths. This shows that the Langston and Trabasso (1999; Langston et al., 1999) model hardly changes the connection strength values. The reason why this is generally the case becomes clear when examining the design of this model.

The Langston and Trabasso Model

The input to the Langston and Trabasso (1999; Langston et al., 1999) model takes the form of a network of nodes corresponding to clauses of a narrative text. The connections between these nodes have strength values that depend on the causal relations within the story. Two clause nodes, *i* and *j*, are *causally connected* if the corresponding story events pass the counterfactual test: If *j* would not have occurred without *i* (all other things being equal), and there is no intervening event caused by *i* and causing *j*, then *i* and *j* are causally connected (Langston & Trabasso, 1999, p. 35).

The initial connection strength value w_{ij} of the connection between any pair of nodes *i* and *j* in the text network equals 7 minus the number of causal connections in the shortest causal path between *i* and *j*, with a minimum of 0 (Langston & Trabasso, 1999, p. 36). This means that all nodes are connected to themselves with the maximum strength of 7 ($w_{ii} = 7$), and that the strength of the connection between causally connected clauses *i* and *j* equals $w_{ij} =$

Table 2
Mean Ranks of Events and Corresponding Average Connection Strengths

Story version	Observed rank	Connection strength: Output	Input
Route			
Explained	1.59	4.49	4.35
Unexplained	2.04	3.70	3.62
Time			
Explained	1.88	3.98	3.84
Unexplained	2.61	3.42	3.32
r^2		.88	.89

Note. Mean rank of Route event in the Route story versions or Time event in the Time story versions (Trabasso & Bartolone, 2003) and average connection strength of the corresponding node, both output (Trabasso & Bartolone, 2003) and input (computed by us from Figures 3 and 4 in Trabasso & Bartolone, 2003).

6. Only those connections with a strength of 6 are shown in the text networks of Trabasso and Bartolone (2003, Figures 1, 3, and 4). Connections between nodes that are not causally connected receive strengths between 0 and 5. All connections are symmetrical, so $w_{ij} = w_{ji}$.

The model processes a text one clause at a time. When a clause is read, its node is added to the text network by connecting it to the previous nodes. After determining the initial connection strengths of these new connections, an integration process takes place in which activation spreads through the network and the connection strength values are updated (Langston & Trabasso, 1999, pp. 39–40).

Activation spreading begins by assigning to each node i a positive activation value a_i. The new clause node has an initial value of 1, and any other node begins with the value that resulted from integrating the previous clause. Next, the two-step activation spreading process, identical to that of the construction–integration model (Kintsch, 1988), is applied repeatedly. In the first step, the activation value of each node is set to the sum of the values of all nodes, weighted by the node's connection strengths. Formally, the new activation of a node i is computed by $\Sigma_j a_j w_{ij}$, where j ranges over all nodes in the network at that moment. In the second step, each activation value is divided by the sum of all activation values, resulting in normalized activations that sum to 1. This process is repeated until the total change in activation values falls below an arbitrarily small threshold value.

After activation has settled, the connection strengths are updated. Each strength is increased by an amount equal to the product of its current value and the activations on both ends of the connection, as expressed by:

$$\Delta w_{ij} = w_{ij}\, a_i\, a_j. \tag{1}$$

These updated connection strengths serve as predictors of empirical data.

According to this algorithm, when the first node enters the model, it necessarily receives all activation because there are no other nodes, resulting in $a_1 = 1$. Because its only connection is the connection to itself, with an initial strength of 7, the strength increase equals $\Delta w_{1,1} = 7$, and the updated strength becomes $w_{1,1} = 14$. Assuming that the second clause is causally connected to the first, the initial strength of the connection between the first two nodes is $w_{1,2} = w_{2,1} = 6$. Of course, the second node is also connected to itself with $w_{2,2} = 7$. Because, at this point, the first node's self-connection strength is larger than the second node's, it receives more activation, and its self-connection strength increases more than that of the second node. This effect is amplified, because in Equation 1 the increase in connection strength is multiplied by the strength itself.

It is not hard to see that no connection strength can catch up with the head start of the first node's self-connection. After processing either network in Trabasso and Bartolone (2003, Figure 3), we found that the Langston and Trabasso (1999; Langston et al., 1999) model yielded as the largest strength the first node's self-connection strength: $w_{1,1} = 1.4 \times 10^7$. The second largest strength was the one between the first two nodes and had a much smaller value of $w_{1,2} = 18.3$. Not only are such results unrealistic, they are also inconsistent with the results in Trabasso and Bartolone (Table 5). We found that those data could not be replicated unless we reduced the head start effect by making all self-connection strengths nonadjustable. In other words, although this is not mentioned anywhere, it seems that Equation 1 is valid only for $i \neq j$.

Even after making this small adjustment, however, there remains a strong head start effect in the Langston and Trabasso (1999; Langston et al., 1999) model: Connections between earlier nodes end up with larger strength values than connections between later nodes. Because connection strengths and activation values are always positive, it is immediately clear from Equation 1 that strengths can never decrease.[1] The longer a connection is in the model, the larger its strength will become, so earlier nodes receive larger connection strengths. This effect reinforces itself, because the rise in connection strengths (Equation 1) increases with larger strengths. Moreover, the nodes that are connected with larger strengths receive more activation, which increases Δw_{ij} even more for these nodes. Later nodes, on the other hand, generally do not receive as much activation because they have smaller connection strengths by the time they enter the model. Therefore, these strengths are hardly increased at all. To summarize, because of the head start effect, all but the first few connections end up with strengths very close to their initial values.

There is a clear head start effect in Trabasso and Bartolone's (2003) simulations. Averaged over their six networks, we found that node number accounted for 14.5% of variance in final connection strengths.[2] Taken together, 92.5% of variance in final connection strengths was explained by initial strengths and node number. In other words, the Langston and Trabasso (1999; Langston et al., 1999) model does not do much more than take the initial connection strengths and increase the strengths of connections between the first few nodes. For all other nodes, the connection strengths after processing the complete network are very close to the initial strengths.

Langston and Trabasso (1999) note that "the general tendency for later sentences to be lower in connection strength leads to underestimation of empirical data" (p. 63). Because the model accomplishes not much more than this undesired head start effect, one cannot expect the empirical data to be predicted better by the model's output than they are by its input. Indeed, Tables 1 and 2, in which we compare the predictive value of the final connection strengths reported by Trabasso and Bartolone (2003) with that of the strengths that result if no updating takes place, confirm this expectation.

Conclusion

Many empirical findings in text comprehension research are accounted for by the causal relations among a story's events. By applying a causal analysis, Trabasso and Bartolone (2003) gave an alternative explanation for experimental data by Kahneman and Tversky (1982) and found support for this alternative explanation

[1] Nevertheless, note that in Table 1, the average output connection strength of the "Crossing" and "Boy" nodes reported by Trabasso and Bartolone (2003) was slightly lower than their average input connection strength computed by us from the networks given by Trabasso and Bartolone (Figure 1). We do not have an explanation for this.

[2] This is the proportion of variance in output strengths w explained by i^{-1}, where i is the number of the clause that brought the connection into the network. Using i^{-1} instead of i reduces the importance of later connections, thereby incorporating the self-reinforcing property of the head start effect.

in a replication of the experiment with an improved design. They were right in pointing to the confounding between normality and explanation in the Kahneman and Tversky study. However, Trabasso and Bartolone did not show that these findings are predicted by the Langston and Trabasso (1999; Langston et al., 1999) model, because the connection strengths this model produces fit the data no better than the network's initial connection strengths. Although this does not affect Trabasso and Bartolone's critique of Kahneman and Tversky's conclusion, it does indicate that the model adds no predictive value. In addition, neither Langston and Trabasso (1999) nor Langston et al. (1999) showed that the updated strengths formed a better predictor of empirical data than did the initial strengths. In short, there seems to be no evidence that applying the Langston and Trabasso model improves the predictions made by the initial connection strengths determined by the length of the causal path between each pair of text clauses. This has been shown to be a consequence of the computational setup of the model, which necessarily leads to a strong head start effect for the first few nodes but no change in connection strengths between later clauses of the text.

References

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York: Cambridge University Press.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review, 95,* 163–182.

Langston, M. C., & Trabasso, T. (1999). Modeling causal integration and availability of information during comprehension of narrative texts. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 29–69). Mahwah, NJ: Erlbaum.

Langston, M. C., Trabasso, T., & Magliano, J. P. (1999). A connectionist model of narrative comprehension. In A. Ram & K. Moorman (Eds.), *Understanding language understanding: Computational models of reading* (pp. 181–226). Cambridge, MA: MIT Press.

Lutz, M. F., & Radvansky, G. A. (1997). The fate of completed goal information in narrative comprehension. *Journal of Memory and Language, 36,* 293–310.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language, 26,* 453–465.

Suh, S., & Trabasso, T. (1993). Inferences during on-line processing: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. *Journal of Memory and Language, 32,* 279–301.

Trabasso, T. (in press). The role of causal reasoning in understanding narratives. In T. Trabasso, J. Sabatini, D. W. Massaro, & R. C. Calfee (Eds.), *From orthography to pedagogy: Essays in honor of Richard L. Venezky.* Mahwah, NJ: Erlbaum.

Trabasso, T., & Bartolone, J. (2003). Story understanding and counterfactual reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 904–923.

Trabasso, T., & Sperry, L. L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language, 24,* 595–611.

Trabasso, T., & Van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language, 24,* 612–630.

Received May 11, 2004
Revision received October 5, 2004
Accepted October 8, 2004 ■