Optimization of a multinomial model for investigating hallucinations and delusions with source monitoring

Todd S. Woodward\textsuperscript{a,b,*}, Mahesh Menon\textsuperscript{c}, Xiangen Hu\textsuperscript{d}, Richard S.E. Keefe\textsuperscript{e}

\textsuperscript{a} Department of Research, Riverview Hospital, Coquitlam, Canada
\textsuperscript{b} Department of Psychology, Simon Fraser University, Burnaby, Canada
\textsuperscript{c} Schizophrenia PET Group, Centre for Addiction & Mental Health, Toronto, Canada
\textsuperscript{d} Department of Psychology, University of Memphis, Memphis, USA
\textsuperscript{e} Department of Psychiatry and Behavioral Sciences, Duke University, Durham, USA

Received 2 February 2006; accepted 3 March 2006
Available online 19 April 2006

Abstract

Studies of source monitoring have played an important role in cognitive investigations of the inner/outer confusions that characterize hallucinations and delusions in schizophrenia, and multinomial modelling is a statistical/cognitive modelling technique that provides a powerful method for analyzing source monitoring data. The purpose of the current work is to describe how multinomial models can be optimized to answer direct questions about hallucinations and delusions in schizophrenia research. To demonstrate this, we present a reanalysis of previously published source monitoring data, comparing a group of patients with schneiderian first rank symptoms to a group without schneiderian first rank symptoms. The main findings of this analysis were (1) impaired recognition of self-generated items and (2) evidence that impaired source discrimination of perceived items is accompanied by an internalization bias in the target symptom group. Statistical and cognitive interpretations of the findings are discussed.

\textsuperscript{©} 2006 Elsevier B.V. All rights reserved.

Keywords: Schizophrenia; Source monitoring; Multinomial modelling; Hallucinations; Delusions

Fundamental to an understanding of reality distortion in schizophrenia (i.e., hallucinations and delusions) is the study of inner/outer confusions in cognition (Fowler, 2000). This relationship is most apparent when considering auditory hallucinations, as these can be conceptualized as the assignment of internally generated mental events to an external source. Originally referred to as reality monitoring (Johnson et al., 1993), the study of inner/outer confusions in memory has recently been employed to investigate the cognitive underpinnings of schizophrenia, typically under the rubric of source monitor-
One of the limitations of many of these studies is that guessing strategies (e.g., strategically increasing external-source guesses) have not been separated from memory errors; however, multinomial modelling allows these processes to be separately measured. Multinomial models attempt to explain discrete responses in a particular psychological paradigm by postulating latent cognitive processes that combine in different ways to determine the response category. The basic idea is that any given response category may occur as a consequence of one or more processing sequences, where each processing sequence is characterized by a series of successful or unsuccessful processing events. The processing sequences are represented in a tree structure (see Fig. 1). The root (or initial node) represents the beginning of the processing sequence, the intermediate nodes represent stages involving a choice between two or more processing events, and the terminal nodes correspond to the observable response categories. The application of multinomial models to source monitoring has been reviewed in detail elsewhere (Batchelder and Riefer, 1999).

Multinomial modelling provides separate source misattribution parameters for when items are recognized and when items are not recognized; we refer to the former as biases, and the latter as guesses. It is the biases, which affect reactivated “in mind” cognitive representations (i.e., recognized items in the source monitoring memory context), that are of primary interest for the study of hallucinations and delusions (Keefe et al., 2002, p. 53). Multinomial models used in previous studies of schizophrenia (Keefe et al., 2002, 1999) were developed within the context of cognitive psychology, but were not optimized for the study of hallucinations and delusions. In traditional multinomial models, parameters are created for old and new recognition (parameter D in Fig. 1), source recognition (parameter d in Fig. 1) and biases. However, estimates of true internalization and externalization biases are not possible, because parameter estimates do not vary with the originating source. That is to say, unlike the recognition (parameters $D_1$...
3, see Fig. 1) and source discrimination (parameters $d_{1-3}$, see Fig. 1) estimates, biases affecting recognized but not discriminated items were equated for items with internal and external origins (referred to as as, ap and ah in Keefe et al., 2002, Fig. 3).

To understand the importance of this distinction, consider that the externalizing account of hallucinations applies specifically to internally generated cognitive operations. Therefore, to properly test the externalization account of hallucinations, a bias towards the external source must be specific to items originating from the internal source, and must not generalize to other items. Multinomial modelling is flexible with respect to the models that can be tested, and externalization and internalization biases for recognized items can be measured for external and internal sources separately, such that an appropriate test for the cognitive underpinnings of reality distortion can be achieved.

In order to directly compare the models developed within the tradition of cognitive psychology with those optimized for schizophrenia research, we reanalyzed previously published data (Keefe et al., 2002). These data are derived from a three-source monitoring experiment, for which 16 words are presented from each of three sources: self (word stem completed by subject), picture (pictures presented on a card) and heard (words read aloud by the experimenter). In a recognition trial, these 42 previously presented words, along with 16 new words, were presented on a test sheet, and subjects were asked to make new/old and source judgements at their own pace. Two groups of patients split on the presence of schneiderian first rank symptoms were compared; the originally published patient–control comparison was not carried out using the model introduced here (explained in more detail below). Sample sizes are presented in Table 1, but for all other aspects of sample description and methodology, please refer to the Methods section of the original article (Keefe et al., 2002).

### Table 1
Response frequencies for target vs. non-target schizophrenia patients as a function of source and response (from Keefe et al., 2002)

<table>
<thead>
<tr>
<th>Source</th>
<th>Response</th>
<th>Self-generated</th>
<th>Picture</th>
<th>Heard</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target symptom group (N = 18)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-generated</td>
<td></td>
<td>323 (56)</td>
<td>55 (10)</td>
<td>95 (16)</td>
<td>103 (18)</td>
</tr>
<tr>
<td>Picture</td>
<td></td>
<td>21 (4)</td>
<td>502 (87)</td>
<td>19 (3)</td>
<td>34 (6)</td>
</tr>
<tr>
<td>Heard</td>
<td></td>
<td>57 (10)</td>
<td>35 (6)</td>
<td>239 (41)</td>
<td>245 (43)</td>
</tr>
<tr>
<td>New</td>
<td></td>
<td>32 (6)</td>
<td>25 (4)</td>
<td>90 (16)</td>
<td>429 (74)</td>
</tr>
<tr>
<td><strong>Non-target symptom group (N = 11)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-generated</td>
<td></td>
<td>234 (66)</td>
<td>30 (9)</td>
<td>42 (12)</td>
<td>46 (13)</td>
</tr>
<tr>
<td>Picture</td>
<td></td>
<td>3 (1)</td>
<td>322 (91)</td>
<td>6 (2)</td>
<td>21 (6)</td>
</tr>
<tr>
<td>Heard</td>
<td></td>
<td>17 (5)</td>
<td>22 (6)</td>
<td>164 (47)</td>
<td>149 (42)</td>
</tr>
<tr>
<td>New</td>
<td></td>
<td>18 (5)</td>
<td>16 (5)</td>
<td>52 (15)</td>
<td>266 (76)</td>
</tr>
</tbody>
</table>

Row percentages are presented in parentheses. Correct responses are set in bold font.

1 The abscissas of the originally published (Keefe et al., 2002) Fig. 3 and Fig. 4 are mislabeled: they refer to categories of items, but should refer to categories of responses.

### 1. Modelling methodology

#### 1.1. Model development

Existing multinomial models (e.g., Bayen et al., 1996; Keefe et al., 2002, 1999; e.g., Riefer et al., 1994) are not ideal for schizophrenia research on source monitoring for two reasons, both of which are concerned with our theoretical stance on the processing of recognized but not discriminated items in schizophrenia. First, true internalization and externalization biases that are appropriate candidates for the cognitive underpinnings of reality distortion cannot be separately estimated using existing models. That is to say, in the existing model, the probability of biasing responses towards an internal or external source does not vary with the source of the information. If hallucinations involve transposing internal source information into external source information, to detect this, estimation of source-specific biases are necessary. Second, an estimate for confusion between the two external sources for recognized items was not included in existing models. This is important because the symptoms associated with externalization and internalization biases (internal–external confusion) should not extend to confusion of external sources (external–external confusion).

An Excel spreadsheet specialized for multinomial modelling (Dodson et al., 1998) was modified to
carry out the analysis reported below. For all tests of the statistical/cognitive models, starting values for probability estimates were 0.5, and $G^2$ was used to assess overall fit. The log likelihood ratio statistic $G^2$ asymptotically has a $\chi^2$ distribution (Riefer and Batchelder, 1988). The final model is displayed in Fig. 1. The fundamental steps are described in detail elsewhere (Woodward and Menon, 2005).

1.2. Model parameters and assumptions

The assumptions underlying the statistical/cognitive model used in this research diverged from that traditionally used (e.g., Bayen et al., 1996; Keefe et al., 2002, 1999; e.g., Riefer et al., 1994), so must be made explicit. Traditional models assume that biases occur independently from the originating source. In contrast, we assume that the contextual information that is carried with “in mind” (i.e., recognized as old) memory traces consists of three components: (1) the signal reflecting the correct source-dependent information (e.g., perceptual vividness for external sources, and generating thoughts for internal sources). The probability of the signal determining the response is captured in $d$ parameters; (2) the source-dependent biases reflecting the misleading source information (e.g., perception-elicited thoughts for external sources, imagined perceptual vividness for internal sources). The probability of the biases determining the response is captured in $a$ and $e$ parameters, and (3) noise that contains information not informative with respect to the response of interest (e.g., what the experimenter was wearing; this is not captured by the model). According to this model, the primary determinant of any given response depends on the strength of both the signal and the biases. The probability of biases dominating the signal, resulting in a source misattribution, is reflected by the $a_1$, $a_2$ and $a_3$ parameters (externalization, internalization for picture item, and internalization for heard items, respectively), and by the $e_1$, $e_2$ and $e_3$ parameters (for the specific nature of an externalization, confusion of external sources for picture items, and confusion of external sources for heard items, respectively). The biases may determine the response if the signal is not detected for a given trial, resulting in an error. If neither the signal nor the biases determine the response, the correct internal/external decision may still be reached based only on the information that allowed correct recognition, and this is reflected by the $1 - a_1 \rightarrow 1 - e_1$, $1 - a_2 \rightarrow 1 - e_2$ and $1 - a_3 \rightarrow 1 - e_3$ processing tree branches.

As is presented elsewhere in more detail (Woodward and Menon, 2005), the $d$ and $a$ parameters can offset one another to maintain fit when statistically comparing across symptom groups. For example, an artificial reduction in $a$, and the associated reduction in the estimated influence of bias, necessitates an increased probability of correct responses by way of the $1 - a_1 \rightarrow 1 - e_1$ and $1 - a_2 \rightarrow 1 - e_2$ routes in order to maintain fit with the response frequencies. In this situation, the model will compensate by decreasing the $d$ route for correct responses, and typically no reduction in fit will be observed. For these reasons, our statistical tests with respect to source discrimination and biases must be stated in terms of both $a$ and $d$ parameters, as opposed to individual parameters, because substantial fluctuations in the $d$ and $a$ parameters, respectively, can be induced by the model to compensate for poor fits with respect to the $a$ and $d$ parameters, respectively. A between-group difference in the $\{d, a\}$ parameter set suggest either (1) a combined impairment, involving impairment in discriminating the source and an increased influence of bias, or (2) no impairment at discriminating the source, but a severe bias, or finally (3) a severe impairment at discriminating the source, but no bias. This cognitive/statistical model does not provide a method of distinguishing between these three possibilities; however, visual inspection of the parameter estimates can typically reasonably rule out one or two of these three possibilities (see below).

Previous research using multinomial modelling suggests that guessing parameters (Keefe et al., 2002; Woodward et al., 2005) and confusion of external sources parameters (Woodward et al., 2005) do not differ between groups of patients split on positive symptoms. Therefore, in the development of the model presented in Fig. 1, the parameters $e_1$, $e_2$, $e_3$, $b$, $g$, and $g_1$ were equated across symptom groups; this saturated model cannot be statistically tested, but the $G^2$ value was very low ($G^2=0.70$). The values for the parameters that were equated across groups in the base (fully specified) model were as follows: $e_1=0.64$, $e_2=0.64$, $e_3=0.58$, $b=0.25$, $g=0.79$, and $g_1=0.78$. The $e_2$ and
e₃ values indicate that confusion of external sources is comparable for heard and picture; that is to say, for approximately 60% of recognized external items for which the source signal is not detected, incorrect external source bias determined the response. The e₁ value indicates that, for an estimated 64% of recognized but not discriminated self-generated items, bias leads the subject to respond incorrectly with the heard source. The b, g, and g₁ values correspond to old, external and heard guesses, respectively, and correspond very closely to the proportion of old, external and heard responses given for new items (see the “new” rows in Table 1).

The comparison of controls to patients could not be carried out using this model. This is because due to the considerable patient–control discrepancy in most cognitive abilities (Heinrichs and Zakzanis, 1998), very few parameters could be equated across groups while maintaining model fit, requiring too many degrees of freedom, and leading to an over-specified model. However, elsewhere (Woodward and Menon, 2005) we show how a model optimized for the study of hallucination and delusions can be used to compare patients to controls using a different cognitive paradigm. Therefore, for the current analysis, we focussed on comparing the two patients groups (target vs. non-target symptoms).

1.3. Group comparison methodology

Based on Fig. 1, we carried out between-group comparisons, using previously published methodology for group comparisons (Keefe et al., 2002, p. 65). For between-group comparisons, if the difference between target and non-target symptom groups with respect to parameter D₁ (for example) was of interest, a model was devised that included both target and non-target symptom groups, and had a total of 24 parameters that were free to vary between groups. There are 9 parameters (D₁, D₂, D₃, d₁, d₂, d₃, a₁, a₂, a₃) for each group (target and non-target symptom groups) and 6 parameters (e₁, e₂, e₃, b, g, g₁) shared by the two groups. This gives a total of 24 distinct parameters.

After the model was established, we conducted a goodness-of-fit test following common methods of testing nested hypothesis (Batchelder and Riefer, 1990), where any restricted model represents H₀, and any non-restricted model H₁. For example, the first hypothesis test uses the original model, namely, the model with all 24 parameters as H₁:

\[ H₀, \quad D₁ \text{ (target symptoms)} = D₁ \text{ (non-target symptoms). Total number of parameter is 23.} \]

\[ H₁, \quad \text{All 24 parameters are free to change. Total number of parameter is 24.} \]

For this significance test, the \( G² \) for the 24-parameter model is subtracted from the \( G² \) for the 23-parameter model. The difference of two \( χ² \) distributed \( G² \) values is also \( χ² \) distributed. The degrees of freedom for the resulting nested hypothesis test is computed as difference between the degrees of freedom for models \( H₁ \) and \( H₀ \). In this case, \( χ² \) test of differences between two groups has 1 degree of freedom, but group differences in sets of parameters can also be tested between groups if the degrees of freedom are adjusted appropriately. Other conclusions can be drawn by carrying out similar types of analyses: identify \( H₀ \) and \( H₁ \), where \( H₀ \) is a submodel of \( H₁ \).

2. Results

The originally published response frequency table is presented in Table 1, with row percentages added. The parameter estimates for the contrast of schizophrenia patients with target symptoms and those without target symptoms are displayed in Table 2. Recognition parameters differed significantly between the groups for self-generated (\( D₁ \)) items only, \( G²(1)=4.19, \ p<0.05 \). When observing the sets of source discrimination parameters, significant between-group differences were observed for both picture \( \{d₂, a₂\} \) and heard \( \{d₃, a₃\} \) internalization sets, \( G²(2)=10.03, \ p<0.01, \)

Table 2

<table>
<thead>
<tr>
<th>Group</th>
<th>D₁</th>
<th>D₂</th>
<th>D₃</th>
<th>d₁</th>
<th>d₂</th>
<th>d₃</th>
<th>a₁</th>
<th>a₂</th>
<th>a₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>0.76⁺</td>
<td>0.92</td>
<td>0.43</td>
<td>0.47</td>
<td>0.93⁺</td>
<td>0.71⁺</td>
<td>0.53</td>
<td>0.50⁺</td>
<td>0.55⁺</td>
</tr>
<tr>
<td>Non-target</td>
<td>0.83</td>
<td>0.92</td>
<td>0.43</td>
<td>0.44</td>
<td>0.99</td>
<td>0.82</td>
<td>0.37</td>
<td>0.36</td>
<td>0.23</td>
</tr>
</tbody>
</table>

See Fig. 1 caption for parameter definitions.

⁺ \( p<.01 \) when parameters were simultaneously constrained to be equated between target and non-target symptom groups.

⁺⁺ \( p<.01 \) when parameters were simultaneously constrained to be equated between target and non-target symptom groups.

⁺⁺⁺ \( p<.05 \) for target symptoms vs. non-target symptom groups.
$G^2(2) = 9.56$, $p < 0.01$, respectively, but not for the externalization set $\{d_1, a_1\}$, $G^2(2) = 4.35$, $p = 0.11$.

3. General discussion

The purpose of the current work was to describe how multinomial models can be optimized to answer direct questions about hallucinations and delusions in source monitoring investigations of schizophrenia. To demonstrate this, we presented a reanalysis of previously published source monitoring data, for which a group of patients with schizophrenia was compared when divided on a set of target symptoms (viz., Schneiderian first rank symptoms). This analysis found the previously reported (Keefe et al., 2002) impaired recognition of self-generated items, and impairment in source discrimination of perceived (heard and picture) items, for the group with Schneiderian first rank symptoms. A novel finding was that this impairment in source discrimination of perceived items was accompanied by an internalization bias.

The original interpretation of impaired recognition of self-generated items for the target symptom group, combined with intact recognition of externally generated stimuli, holds with the current analysis: this is evidence for autonoetic agnosia underlying specific symptoms of schizophrenia, that is to say, the inability to identify self-generated events (Keefe et al., 2002). The impairment in source discrimination of perceived (heard and picture) items for the target symptom group was previously interpreted as an unexpected result (Keefe et al., 2002, p. 61), but the current analysis, and other research using this approach, aids interpretation. Namely, there is evidence for the contribution of an internalization bias to the source discrimination deficit, replicating previous findings (Woodward et al., 2005). We interpret this association between internalizations and delusions as indexing a cognitive marker of the self-referential and/or self focussed nature of delusional thinking (Birchwood, 1999, p. 316; Freeman et al., 2000), and it can readily be incorporated into the inner/outer confusions characteristic of Schneiderian delusions.

Inspection of the frequency table (Table 1) reinforces this conclusion. Comparing row percentages between groups, for self-generated items, the incorrect responses for the target symptom group are distributed fairly evenly over the “picture”, “heard” and “new” responses categories. Contrast this with picture and heard items. For these items, the incorrect responses tend to disproportionately cluster in the “self-generated” response category, but not in the “new” or the incorrect external source response category. The current analysis allows this “internalization” pattern of response frequencies whereas with the previous method of analysis it was missed, because the $a$ parameter was averaged over sources, instead of varying with source (as in the current analysis).

The originally published hypotheses for these data (Keefe et al., 2002) proposed that an “externalization” pattern should be associated with the positive symptoms of schizophrenia. The current analysis of these data, and previous work, suggest that the symptoms on which the patient group is split may determine the directionality of the source monitoring bias. More specifically, in previous work (Woodward et al., 2005), an externalization bias was detectable when patients where a different sample of patients were split on 2nd person auditory hallucinations. In contrast, for the current data, an internalization bias was observed when the patients were split on Schneiderian delusions and/or 3rd person (Schneiderian) auditory hallucinations, replicating the internalization bias observed previously when a different sample of patients were split on delusions (Woodward et al., 2005).

This study supports the contention that the cognitive systems involved in source monitoring underlie the positive symptoms in schizophrenia. Specifically, the results support the postulation that delusional patients display evidence for an internalization bias. An internalization bias presumably works in combination with a number of other cognitive mechanisms to form a delusional belief system (Frith, 1994; Garety et al., 1991; Menon et al., 2005; Moritz and Woodward, 2006; Woodward et al., in press-a,b). Our goal with this demonstration was to encourage schizophrenia researchers to use statistical methodology that allows more precise characterization of error types. Further work may also address more specific hypotheses about the relationship between the internal/external direction of biases and specific symptoms.
Acknowledgement

We thank Shurin Hase for assistance with manuscript preparation.

References


