Paper:

Lowe, R. & Norman, P. (2016). Information processing in illness representation: Implications from an associative-learning framework.. *Health Psychology*
http://dx.doi.org/10.1037/hea0000457

Running head: INFORMATION PROCESSING IN ILLNESS REPRESENTATION

Information processing in illness representation:

Implications from an associative learning framework

Rob Lowe

Swansea University

Paul Norman

University of Sheffield

Direct correspondence to:

Rob Lowe, Psychology Department, Swansea University, Singleton Park, Swansea, SA2 8PP, UK, email:

r.p.lowe@swansea.ac.uk

Abstract

Objective: The common sense model (CSM) outlines how illness representations are important for understanding adjustment to health threats. However, psychological processes giving rise to these representations are little understood. To address this, an associative learning framework was used to model low-level process mechanics of illness representation and coping decision-making. Methods: Associative learning was modeled within a connectionist network simulation. Two types of information were paired: illness identities (indigestion, heart attack, cancer) were paired with illness belief profiles (cause, timeline, consequences, control/cure); and specific illness beliefs were paired with coping procedures (family doctor, emergency services, self-treatment). To emulate past experience, the network was trained with these pairings. As an analogue of a current illness event, the trained network was exposed to partial information (illness identity or select representation beliefs) and its response recorded. Results: The network a) produced the appropriate representation profile (beliefs) for a given illness identity, b) prioritized expected coping procedures and c) highlighted circumstances where activated representation profiles could include self-generated or counter-factual beliefs. Conclusions: Encoding and activation of illness beliefs can occur spontaneously and automatically; conventional questionnaire measurement may be insensitive to these automatic representations. Furthermore, illness representations may comprise a coherent set of non-independent beliefs (a schema) rather than a collective of independent beliefs. Incoming information may generate a 'tipping point', dramatically changing the active schema as a new illness knowledge set is invoked. Finally, automatic activation of well-learned information can lead to the erroneous interpretation of illness events, with implications for (inappropriate) coping efforts.

Keywords: illness representation; common sense model; connectionist; network; memory; associative learning

Information processing in illness representation: Implications from an associative learning framework

The Common Sense Model (CSM) was developed by Leventhal, Meyer and Nerenz (1980) to explain people's reactions to health threats. Central to this is explaining the way individuals form a cognitive representation of the health threat. Leventhal et al. outlined five core dimensions of such representations; identity (the label given to the health threat); cause (beliefs about the origin of the health threat); timeline (beliefs about the duration of the health threat and whether it is acute, cyclical, or chronic); consequences (beliefs about the physical, psychological, social, and economic outcomes of the health threat); and control/cure (beliefs about the extent to which the health threat is amenable to treatment). The construct validity of these dimensions has been confirmed in a number of studies (Bauman, Cameron, Zimmerman & Leventhal, 1989; Lau, Bernard & Hartman, 1989; Meyer, Leventhal & Gutmann, 1985), and are typically measured by the Illness Perception Questionnaire (IPQ) (Weinman, Petrie, Moss-Morris & Horne, 1996), although more recent IPQ versions assess additional dimensions (Broadbent, Petrie, Main, & Weinman, 2006; Moss-Morris et al., 2002). The CSM proposes that individuals' representations of the health threat determine how they will cope. For example, an illness represented as an acute infection may lead a person to take medication to destroy the infection-causing bacteria, whereas its representation as a chronic condition may lead a person to make lifestyle changes, such as exercising, to increase their long-term well-being and survival chances. These are common sense responses, seen as necessary and appropriate, as perceived by the individual (Leventhal, Diefenbach, & Leventhal, 1992).

Meta-analysis (Hagger & Orbell, 2003) confirms associations between illness representations and various ways of coping. Despite this, little attention has been focused on how illness representations are generated. According to Leventhal et al. (1980) illness representations are formed on the basis of the individual's own experiences of the illness (e.g. symptoms), information from others (e.g. family members, doctors) and general "lay" beliefs about their illness. Leventhal, Leventhal and Breland (2011) highlight the importance of prototypes in illness representation and their association with coping procedures; people will have prototypical knowledge about various health threats, their symptoms and what to do about them. Prototypes are stored representations of sets of psychologically differentiable but related targets or concepts (i.e. category members) developed from experience (Fiske & Taylor, 2013). A prototypical illness representation will comprise knowledge – beliefs – about the illness, acquired through experience and social learning. This information enables the individual to form a representation of an encountered health threat (Leventhal et al., 1980). Because beliefs reflect acquired knowledge that is stored and recalled, they are phenomena of memory. However, what are the processing mechanics for activating and combining beliefs into an illness representation? How do these representations trigger coping responses?

We explore these questions by placing the CSM within an associative memory framework. Below we examine illness representations as events in memory and the application of connectionist networks for investigating process mechanics of representation encoding and retrieval. This level of analysis may yield useful insights about why certain representations are formed, whether this process is automatic, whether component beliefs are independent or part of a knowledge set (schema), why representations may include inaccurate or counter-factual beliefs, how illness representations trigger coping responses, and how illness representations and coping may change as symptoms unfold over time. This in turn has implications for the measurement of illness representations, currently dominated by explicit measures such as the IPQ (Broadbent et al., 2006; Moss-Morris et al., 2002; Weinman et al., 1996).

Illness representations as memory events

Consistently co-activated concepts can become linked in memory through associative learning (Eiser, 1994; Smith, 1996, 2009; Van Overwalle & Siebler, 2005). Hence, illness representations may comprise coherent *sets* of linked beliefs (knowledge) associated with specific health threats, symptoms and coping procedures (i.e. a schema) (Lowe, Porter, Snooks, Button & Evans, 2011). For example, when facing symptoms and uncertain of the diagnosis, people may use knowledge from similar experiences to make predictions about what is wrong and how the situation will unfold. Leventhal, Nerenz and Steele

(1984) illustrate this process with an example of 'John' who initially misconstrues his symptoms as indigestion and tries to continue normal activities. Symptoms worsen, and John suspects it may be something more serious (stomach ulcer, liver disease). Eventually, the pain becomes severe, his GP is contacted, and he is admitted to hospital as an emergency and ultimately diagnosed with a mild coronary.

An influential approach for understanding encoding and retrieval of information in associative memory is *connectionism* (Eiser, 1994; Smith, 1996). This has been successfully applied to modeling social-cognitive phenomena such as attitudes, attributions, prejudice, and personality (Eiser, Fazio, Stafford & Prescott, 2003; Eiser, Stafford & Fazio, 2009; Shoda & Mischel, 1998; Van Overwalle, 1998; Van Overwalle & Siebler, 2005; Van Overwalle & Van Rooy, 1998). Connectionism involves the use of *connectionist networks* for understanding the mechanics of information computation that underlies cognition (Callan, 1999; O'Reilly & Munakata, 2000; O'Reilly et al., 2012). Connectionist networks are usually implemented in the virtual environment of a computer simulation (e.g. McLeod, Plunkett & Rolls, 1998; O'Reilly & Munakata, 2000). We used this approach for examining the way illness representations are construed within an associative learning framework.

In constructing the connectionist model, we assumed two sets of information become linked in memory through experience and learning. First, people associate an illness label with beliefs about the illness. Second, specific illness beliefs are associated with particular coping procedures (Leventhal et al., 2011). If this system was cued with an illness label, associated beliefs would be triggered along with their allied coping representations. Indeed, McAndrew et al. (2008) describe a top-down process where a person's knowledge they have a specific illness can activate linked beliefs and coping procedures. Conversely, beliefs associated with symptoms for an initially unknown illness can activate allied beliefs, candidate illness label(s) and coping representations – a bottom-up process seen in acute conditions where symptoms signal the presence of a problem (McAndrew et al., 2008). Thus, an athlete experiencing pain in their knee may attribute it to a poor maneuver (cause), believe it to be a strain (label) that is not serious (consequences) and likely pass quickly (timeline), and expect that it will get better through rest (coping).

Our connectionist model tested this assumed structure for three illnesses: indigestion, heart attack, and cancer. Figure 1a shows associations of illness labels with specific beliefs. For our purpose, indigestion was associated with a fast (acute) timeline, modest consequences, high control, and a lifestyle cause; a heart attack was associated with a fast timeline (unfolds quickly), serious consequences, low control, and a lifestyle cause; cancer was associated with a slow (chronic) timeline, serious consequences, low control, and a dispositional (e.g. genetic) cause. Although these are hypothetical, research shows that, for example, lifestyle and genetics/familial factors are key attributions for heart disease (French, Senior, Weinman & Marteau, 2001) and cancer (Kwate, Thompson, Valdimarsdottir & Bovbjerg, 2005) respectively. Figure 1b shows associations between specific illness beliefs and coping procedures. Here, seeing the GP (family doctor) was associated with both acute and chronic illnesses (short and long timeline) and illnesses with modest or serious consequences; use of emergency services was associated with illnesses that unfold quickly (short timeline) and which are serious; self-treatment was associated with illnesses that have modest consequences and which are controllable (treatable).

We predicted that the connectionist network would successfully associate illness labels with a set of illness beliefs (H_1); the network would associate beliefs with specific coping procedures which could be activated subsequently when presented with an illness label (H_2); partial information about an unidentified illness would activate a wider belief set (but which may not necessarily reflect the reality of the illness situation) (H_3); illness representations and associated coping procedures would change over time as symptoms evolve (H_4).

Method

The connectionist network

The network was simulated using MATLAB 2013b computer software, applying a form of learning called error *backpropagation* (following Callan, 1999). Backpropagation was deemed appropriate because it is an established, powerful, yet parsimonious approach to learning which has successfully modeled belief-based phenomena elsewhere (e.g. Eiser, Stafford & Fazio, 2009; Van Overwalle & Siebler, 2005). The type of backpropagation network was a three-layer *autoencoder* (Rumelhart, Hinton & Williams, 1986). When presented an input pattern, an autoencoder reproduces the same pattern as its output. This system can model content addressable memory, reproducing a full pattern from a noisy or incomplete stimulus. Thus, if presented with the illness label (only) at input, a system trained with an illness-belief pairing should be able to reproduce both the illness label and the belief as its output.

A typical backpropagation architecture was used (Figure 2) in that the network comprised clusters of units (nodes) arranged into three layers. Each unit was fully connected to all units in the next layer up. The network comprised 14 *input* units, 5 *hidden* units, and 14 *output* units. In detail, the first layer was where a stimulus pattern was imposed on the network (the *input layer*). For simplicity, each unit represented a discrete psychological construct. There were 14 input units with one unit representing each of: illness label (indigestion, heart attack, cancer), timeline (fast, slow), consequences (modest, serious), control/cure (low, high), cause (lifestyle, disposition), and coping (GP, emergency services, self-treat). Activation was sent, via weighted connections (the arrows in Figure 2), to the second layer (the *hidden layer*). This layer had fewer units than the input layer (here, five units) which restricted representational capacity and forced the network to extract central relationships within the data and treat less useful information as noise. Essentially, different inputs producing a very similar output were collapsed together into the same hidden layer representation; inputs producing their own particular output were represented by a unique hidden layer representation (McLeod, Plunkett & Rolls, 1998). This enabled the system to evolve *prototypes* to represent critical input-output relationships. [Pilot runs indicated this hidden layer size (five units) was sufficient to accommodate relationships in the data used for our simulations.] Hidden layer units sent activation (again, via weighted connections) to the *output* layer, which was where the system produced its response. Each output unit represented a discrete psychological construct; there were 14 units, mirroring the input layer size so that it could reproduce the input pattern.

Weighted connections are adjustable, which is how the system learns. As such, long-term information is stored on the connections, whereas activation on the units reflects what is on the system's 'mind' at any given moment (O'Reilly & Munakata, 2000). The pattern which the system should produce (here, a reconstruction of the input pattern) is known as the *target*. Learning input-target associations occurs during *training,* which emulates prior experience of co-occurring health threats, symptoms, beliefs, and/or behaviors. During a training event, a stimulus is imposed on the input layer, and activation automatically passes through the weighted connections to ultimately produce a pattern of activation on the output layer. Initially, the system knows nothing about stimulus-target associations, and produces a random response. The network used the *generalized delta rule* for learning. That is, the error between the system's own output and the desired target (which here is the same as the input stimulus) is used for learning. Essentially, some (hidden and output) units are receiving too much activation and others too little. The system corrects for this by adjusting incoming connection weights according to the degree and source of the error. When the same stimulus is imposed next time, the output should shift towards the desired target. Over many training cycles, the network is increasingly able to produce the appropriate response to a given stimulus.

Once trained, the network should be able to reconstruct a complete pattern from a partial (incomplete) input. This is analogous to a person experiencing illness symptoms, and based on prior knowledge (perhaps through a previous experience or via social learning), generates a hypothesis as to the nature of the condition. Alternatively, a person may be able to recall relevant information (beliefs) given a specific illness label. This can be emulated by *testing* the trained network. Partial activation of a pattern on the input layer reflects, for example, an encounter with an exemplar illness (label); the system responds by activating associated information (illness beliefs, coping behavior) as output. By examining input-output responding, it is possible to gain an insight into associative memory processing of illness representation.

The current simulations

Two simulations explored distinct, but associated issues. Simulation 1 examined the illness representation profile and coping response associated with a presented illness label. This represents a situation where a person, believing they have a specific (named) illness, recalls information about that illness and what to do about it. Simulation 2 explored what would happen if a person did not know what was wrong, and how their illness representation and coping may change over time as the interpretation of symptoms evolved. Here, select beliefs (presumed to be associated with interpreting specific symptoms) were cued at input, and the network's response, in terms of illness representation and coping, was examined. For both simulations, findings comprised the mean of outputs across 200 simulator runs.

Training and testing stimuli

A null output from a connectionist network may reflect either the non-existence of a relationship in the data, or a failure of the processing system to detect a relationship that exists. Therefore, networks are usually exposed to data containing known associations; a null output would indicate a problem with the network. This is conventionally achieved using artificially generated data whose associations are known in advance (e.g. Eiser, Fazio, Stafford & Prescott, 2003; Eiser, Stafford & Fazio, 2009; Van Overwalle & Siebler, 2005). A similar approach is found in tests of statistical procedures (e.g. Clatworthy, Hankins, Buick, Weinman, & Horne, 2007). Table 1 shows training patterns used herein. Data consisted of two types of information. The upper table section shows how illness labels were paired with beliefs about specific illness features (*illness-labels-to-beliefs*). The lower table section shows how illness beliefs were paired with various coping procedures (*illness-beliefs-to-coping*). Some associations were anticipated to be stronger (better learned) than others. To reflect this, the amount of training between patterns varied; Table 1 shows the relative proportions of item pairings. The intention was that some pairings would be more established in memory and so be influential in constructing a response.

Note that our aim was not to examine a particular finding from the literature. Rather, we examined the more abstract question of how an information processing system might 'compute' available knowledge (whatever that knowledge might be) to construe an illness representation. That is, how beliefs might be structured to solve logical problems posed by implementing a core element of the SCM within an information processing system. For example, as highlighted above, part of the solution was to pair illness labels with beliefs, and beliefs with coping; our implementation never explicitly paired illness labels with coping. Thus, although individual belief labels were to some extent arbitrary, it was important for labels to characterize logical patterns to facilitate interpreting the network's behavior within the context of illness representation. We accept this is a simplification and that people will have a range of beliefs about a given illness. Our point is that, in as much as a person's beliefs are consistent, then the network can explore how such knowledge might be internally structured and predictive.

Prototype patterns comprised strings of numbers; a feature's presence or absence was represented with a '1' or '0' respectively. During stimulus presentation for both training and testing the network (see below), noise was added to the selected case to convert the prototype pattern into an exemplar. The 1 and 0 values were adjusted down or up respectively by a random value of up to 0.2.

Network training. Starting (untrained) network weights were initialized to a random value between 0 and 0.2. We used *interleaved training*; illness-labels-to-beliefs and illness-beliefs-to-coping patterns were combined into a single data set from which items were randomly selected (without replacement) rather than presented as two separate data sets. The network went through the training data set until all patterns had been presented. Each presentation of the whole set is called an *epoch*; the network was trained across 300 epochs. [Pilot runs using a cross validation process (*early stopping*) showed the system achieved a maximal train-test data fit at approximately 550 epochs, but improvement after about 300 epochs was small.] To prevent unstable and suboptimal training, the network was i) constrained in how much it learned at any given moment (*learning rate*) and ii) a proportion of learning in one data presentation was carried over to the next presentation (*momentum*). Following piloting, the leaning rate was set to 0.01 and momentum to 0.8 (the proportion of potential learning implemented as weight change and carried over respectively). A logistic function constrained hidden and output unit activation to between 0 and 1, preventing unrestrained (run-away) escalation. Weights were updated following each pattern presentation.

Network testing. Following training, the network was tested. Here, new or fragmented patterns were presented to see how the network interpreted them based on its prior learning. For simulation 1, testing comprised presenting illness labels – one at a time – by activating the input unit representing a given illness identity. Because the network specialized in pattern completion, it would activate outputs associated (both directly and indirectly) with a given illness label. That is, it would reconstruct a set of illness beliefs associated with that identity (i.e. the illness representation) and allied coping procedures. In simulation 2, select beliefs (but not the illness identity label) were activated at input to explore the construction of a representation for an unidentified illness unfolding over time. Here, consecutive test patterns reflected different phases of an unfolding illness of unknown identity. This aimed to determine how changing beliefs (reflecting changes in symptom interpretation) influences an overall illness representation and coping procedures. Activation on the output units was recorded for subsequent analysis.

For simplicity, illness identity features had two possible ‘states’. For example, timeline could be either ‘short’ or ‘long’, whilst consequences could be either ‘modest’ or ‘serious’. We eased interpretation of the network’s output by collapsing these binary categories into a single summary score for each illness belief. This was done by simply subtracting one output of the pair from the other. Thus, in Figures 3 and 5, a positive (negative) value represents relative belief in a long (short) timeline, severe (modest) consequences, high (low) control/cure, and a lifestyle/behavioral (internal/dispositional) cause.

Results

Simulation 1

Illness representation profiles associated with a presented illness label.

This simulation examined beliefs and coping procedures activated when presented a specific illness label. When presented each illness label (one at a time), the trained network produced an output profile as summarized in Figure 3. Columns represent mean activations (with *SD* bars) across 200 simulator runs. [Exploratory runs showed the simulator produced column patterns similar to those in Figures 3 and 4 within approximately 10 runs that were maintained at 500 runs, indicating quick acquisition of a steady state.] We predicted that illnesses would differ in their representation (belief) profile. This was tested with a two-way ANOVA: illness label (indigestion, cancer, heart attack) X representation dimension (timeline, consequences, control/cure, cause) with repeated measures on both factors. Results confirmed that representation profiles differed as a function of illness label; there was a significant two-way interaction between illness label and representation dimensions (F (6, 194) = 834.02, p < .001, η_p^2 =.96). Post hoc *t*-tests (Bonferroni corrected) compared each illness belief for indigestion with their equivalents for both cancer and heart attack, and also compared cancer beliefs with their heart attack counterparts. All comparisons differed significantly [except the consequences comparison between cancer and heart attack (t_{199} = 1.53, p = .13)]: timeline lowest t_{199} = 7.42, p < .001; consequences lowest t_{199} = 44.87, p < .001; control lowest t_{199} = 8.72, p < .001; cause lowest t_{199} = 15.49, p < .001. Profiles followed the anticipated pattern: indigestion was construed as having a short timeline, modest consequences, relatively controllable, and having a lifestyle cause. Cancer was regarded as unfolding more slowly, serious, relatively uncontrollable, and having an internal/dispositional cause. A heart attack was seen as something unfolding quickly, serious, relatively uncontrollable, and having a lifestyle cause. In sum, the learning system successfully associated illness labels with a specific profile of illness beliefs. Profiles described characteristic knowledge about each illness.

Coping associated with a presented illness label.

During training, specific illness beliefs were paired with particular coping procedures. Figure 4 shows the mean activation (and *SD* bars) of each coping procedure for each illness. Data were analyzed by two-way ANOVA: coping procedures (attend GP, emergency services, self-treat) X illness label (indigestion, heart attack, cancer) with repeated measures on both factors. Findings confirmed our predictions; the network prioritized seeing a GP for cancer, emergency services for a heart attack, and self-treatment for indigestion. That is, there was a significant interaction between coping procedures and illness label (F (4, 196) = 475.76, p < .001, η_p^2 =.91). Post hoc *t*-tests (Bonferroni corrected) showed that activation of the GP coping response was significantly higher for cancer than when presented with either indigestion (t_{199} = 26.45, p < .001) or heart attack (t_{199} = 21.60, p < .001) illness labels. Activation of accessing emergency services was significantly higher for heart attack than when presented with indigestion (t_{199} = 11.31, p < .001) or cancer (t_{199} = 24.55, p < .001) illness labels. Finally, activation of the self-treatment response was significantly higher for indigestion than when presented with either a heart attack (t_{199} = 34.46, p < .001) or cancer (t_{199} = 7.51, p < .001).

Simulation 1 demonstrated how an associative memory system could learn and retrieve illness-specific beliefs, and use this information to predictably prioritize a coping response. However, we also wanted to explore how a person may try to 'make sense' of symptoms for an unknown illness as it unfolded over time. Simulation 2 looked at how the associative learning system would respond when presented changing illness beliefs in the absence of an illness label.

Simulation 2

Illness representation profiles activated for an unknown illness as it unfolds over time.

When people feel unwell, but have no label for the problem, beliefs associated with presenting symptoms will be used to construct a representation of the situation at that point in time. This information is used to determine how to cope, with prioritized strategies changing as the situation evolves (Leventhal, Nerenz & Steele, 1984). We simulated an unfolding illness, looking at representations activated from incomplete illness information, and how changing representation profiles influenced coping. Three issues were of interest. First, whether the system could use incomplete illness information (symptom-related beliefs) to construe a fuller illness representation (belief profile) based on existing knowledge. Second, whether this self-reconstruction contained beliefs inconsistent with information from the presenting situation. In other words, whether activated knowledge would contain counter-factual beliefs. Finally, how the changing illness representation influenced coping as symptom interpretations evolved.

The network was trained as described for Simulation 1 above. However, testing involved presenting separate patterns, one after the other, each representing a different 'phase' of an illness. Specifically, the network was sequentially presented three incomplete – but somewhat different – illness belief profiles reflecting an unfolding understanding of the situation over three time-points. An illness label was not presented (its identity being unknown). The first belief profile was for *something transient* (analogous to Leventhal and colleagues' (1984) example of John believing he had indigestion), which comprised: modest consequences, low/uncertain control, and a lifestyle/behavioral cause. Note that attributing cause is associated with the experience of illness symptoms (Leventhal, Leventhal & Breland, 2011). As such, a causal inference was included despite the pattern representing an illness whose specific label was unknown. The second symptom interpretation was for *something serious*, and comprised beliefs about serious consequences and low/uncertain control. The final symptom interpretation was for *something urgent*, and consisted of beliefs about a short timeline, serious consequences, and low/uncertain control. The shift from something transient to something serious involved changing consequence beliefs from *modest* to *serious* and causal beliefs from *lifestyle/behavior* to *unknown*. The change from something serious to something urgent involved the addition of one belief (*fast/short* timeline) to the existing profile.

Four observations were of relevance. Firstly, as symptoms evolve over time, the associated illness representation should also change, with knock-on effects for coping choice. Thus, we expected the activated illness representation profiles as well as prioritized coping to differ across illness time points as the symptom interpretation evolved. Mean activation (and *SD*) of illness representation profiles for the three phases of the unfolding situation are shown in Figure 5. Columns indicate activated beliefs at output, and arrows at the top indicate the direction of beliefs cued at input. The difference (change) in belief profiles over the three time phases was assessed with a two-way ANOVA: time phase (something transient, something serious, something urgent) X representation profile (timeline, consequences, control/cure, cause) with repeated measures on both factors. Representation profiles significantly differed according to time phase. That is, there was a two-way interaction between illness representation and time phase (F (6, 194) = 2548.74, $p < .001$, η_p^2 =.99). The pattern of coping across the three illness phases is shown in Figure 6. A two-way ANOVA compared coping (GP, call emergency services, self-treat) across illness time phase (something transient, something serious, something urgent) with repeated measures on both factors. Findings showed that coping varied according to illness phase (F (4, 196) = 2069.01, $p < .001$ η_p^2 =.98). Post hoc t-tests (Bonferroni corrected) showed that for something transient, activation of self-treatment was significantly higher than either seeing the GP (t_{199} = 27.76, $p < .001$) or calling emergency services (t_{199} = 21.05, $p < .001$). As the situation evolved into something serious, coping priority shifted to seeing the GP, which was significantly higher than either calling emergency services (t_{199} = 15.65, $p < .001$) or self-treatment (t_{199} = 69.15, $p < .001$). When the illness was finally deemed urgent, calling emergency services became the priority over seeing the GP (t_{199} = 22.16, $p < .001$) and self-treatment (t_{199} = 99.53, $p < .001$). Overall, the pattern of coping changed as the illness representation evolved over time.

The second observation was activation of those illness beliefs which were not also presented as input stimuli (the columns in Figure 5 without an arrow above them: timeline for something transient and something serious, and cause for something serious and something urgent). Despite the absence of situational information (input activation), the network inferred these illness beliefs (their activation significantly differed from zero; lowest t_{199} = 9.86, $p < .001$). This shows how people may use existing knowledge to generate hypotheses or expectations about an illness despite an absence of information.

Third, when facing an unknown illness, the system invoked *profiles* for (relatively) closely matching known conditions, rather than piecemeal activation of independent beliefs. Correlations comparing the patterns of belief activation (columns) shown in Figure 5 with those in Simulation 1 involving a known illness (Figure 3) reveals they are similar: something transient with indigestion (r = .89, $p < .001$); something serious with cancer (r = .80, $p < .001$); something urgent with heart attack (r = .71, $p < .001$). In terms of activated illness label (not shown), for something transient, the highest label activation was indigestion – higher than either heart attack or cancer (lowest t_{199} = 21.22, $p < .001$); for something serious, the cancer label had the highest activation (lowest t_{199} = 32.61, $p < .001$), and for something urgent, the heart attack label had the highest activation (lowest t_{199} = 24.25, $p < .001$). The network used information in memory to infer a representation profile for the current (unfamiliar) illness and generate hypotheses about identity.

The final observation was activation of the control/cure belief for something transient. The system had been explicitly cued with input information indicating the situation was not controllable, yet the network responded with an internally generated belief that the situation was controllable. In Figure 5, the situational cue (the input activation) for this is shown as a white downward (negative) arrow and the network's own activated control/cure belief is shown as a rising (positive) column. The latter (network generated) activation differed significantly from zero (t_{199} = 26.77, $p < .001$) but was in the opposite direction to the situational cue; i.e. the system interpreted the situation as controllable despite an explicit cue that it was not controllable. The system's use of existing knowledge (about indigestion) to generate a counter-factual belief is analogous to a person holding an illness belief which runs counter to information presented by the situation. Thus, automatic activation of well-learned information may lead to an illness being erroneously interpreted with implications for inappropriate coping (in the current simulation, ineffectual self-treatment).

Discussion

Simulation 1 showed how illness representations can be accommodated within an associative learning framework. The network successfully represented known associations between illness labels and illness beliefs (H_1), and between illness beliefs and coping such that activation of the illness label triggered an appropriate coping procedure (H_2). Simulation 2 showed how illness representations and coping evolved as symptom interpretations changed over time (H_4), highlighting speculative and inappropriate illness representation beliefs (H_3) with implications for the adaptiveness of associated coping procedures.

Four general implications for the CSM emerged. First, encoding and activation of illness beliefs can occur spontaneously and automatically. The connectionist network autonomously learned consistencies between illness identities, representation beliefs and coping procedures – knowledge subsequently activated automatically by a stimulus prompt. No executive control was required. This has implications for understanding the way people construe illness representations during their daily lives. Rather than being constructed deliberatively and rationally, they may be reflexive, generated quickly with few resource requirements. As such, people's initial coping responses may be spontaneous, directed by automatically construed representations (e.g. Fazio, 2001) and arise without consideration of alternatives. This automatic construal is consistent with 'if-then' heuristics proposed by Leventhal as a means by which people identify ways of coping from an illness representation (Leventhal, Leventhal & Contrada, 1998). It is unclear the extent to which conventional questionnaires such as the IPQ (Weinman et al., 1996) are sensitive to automatic illness representations because completion requires deliberative introspection. Our evidence emphasizes the need to consider alternative tools for measuring low-level illness representations (Bargh & Chartrand, 2000). For example, Henderson and colleagues used response latency to demonstrate automatic activation of both illness representations and associated coping procedures (Henderson, Hagger & Orbell, 2007; Henderson, Orbell & Hagger, 2009).

The second implication relates to the schematic nature of illness representations – i.e. as a coherent *set* of beliefs. Research has highlighted associations between specific CSM beliefs with illness outcomes (e.g. McSharry, Moss-Morris & Kendrick, 2011). Whilst some analyses treat beliefs as independent (e.g. as predictors in regression), others treat them as sets (e.g. using cluster analysis, or combining items into a single scale) (Bean, Cundy & Petrie, 2007; Clatworthy, Hankins, Buick, Weinman, & Horne, 2007; Lowe, Porter, Snooks, Button & Evans, 2011). Simulation 2 provided support for the latter conceptualization. Information was linked by common association and activated as a schematic set. This occurred even when evidence from the environment contradicted specific individual beliefs within the schema. The potential for self-generated and counterfactual beliefs within illness representations highlights research opportunities for exploring their manifestation and influence on coping. Simulation 2 also highlighted how representation schemas change as an illness unfolds. For example, adding a single piece of central information led to a tipping-point that resulted in a quite different representation profile. Specifically, 'something serious' was associated with information concerning serious consequences and low or uncertain control/cure. The addition of a brief/short timeline shunted the schema to one representing 'something urgent' (second and third representation profiles in Figures 5). This also influenced coping, moving the priority from seeing the GP to using emergency services (Figure 6). A tipping-point hypothesis suggests future experiments into the way changing information influences respondents' representations.

The third implication concerns the way connectionist models can enhance theory (O'Reilly & Munakata, 2000). First, connectionist modeling requires assumptions to be unequivocal. By contrast, many 'box and arrow' models are descriptive and vague about process mechanisms. Here, we explicitly demonstrated process mechanics of automatic illness belief storage and retrieval. Second, the approach helps generate new hypotheses concerning relationships between construct variables. We conjectured two sets of pairings: illness labels with illness beliefs, and illness beliefs with coping. Illness labels were not directly paired with coping procedures (Figure 1); the network automatically inferred this association in simulation 1. Finally, constructing the simulation emphasized the subtle and discriminating nature of illness representations. For clarity and parsimony, binary classifications described illness representation components. Real-world representations would be more nuanced. For example, timeline (modeled here as unfolding quickly or slowly) may differentiate situations advancing in minutes (choking), hours (giving birth), days (infection), weeks (severe sprains), months (broken bones), and years (chronic conditions). A person could acquire intricate illness knowledge throughout a lifetime, and without attempts to explicitly model these cognitive processes, their subtlety may remain unconsidered or underestimated.

Finally, the CSM differentiates between concrete (bottom-up) representations where illness symptoms signal an acute problem, and conditions whose symptoms are ambiguous (silent or chronic) (e.g. Horowitz, Rein & Leventhal, 2004; McAndrew et al., 2008) where effective coping requires a knowledge-based (top-down) representation. For example, someone knowing they have asthma may adhere with preventative medication despite the absence of symptoms. Herein, both top-down and bottom-up representations were accommodated within the same processing structure. Simulation 1 involved a top-down representation where an illness knowledge label activated specific beliefs about that illness, along with allied coping representations. Simulation 2 involved a bottom-up process of changing beliefs activated by an ambiguous evolving condition, triggering various candidate illness identities and coping options across time. Further work is needed for in-depth exploration of variability in symptom prototypicality, and how this influences bottom-up construal of (in)appropriate illness representation and coping.

Notwithstanding the above advantages, the current study has limitations. First, as a conceptual exploration, the CSM was modeled under optimal data conditions. Real-world phenomena are likely noisier and less clear-cut. Our purpose was to test the ability of a general model of associative memory to produce the sorts of behavior predicted by the CSM. As described earlier, examining an information processing *system* required knowing in advance associations extant in the data and seeing if the system could detect these relationships. As highlighted by Eiser, Fazio,Stafford and Prescott (2003) "The contribution of … connectionist simulations is that we can examine learning… in a much purer fashion, that is, unconfounded by any other tendencies that humans might bring to bear on this situation" (p1225). Second, we made assumptions about the way information was paired: illness labels with illness beliefs, beliefs with coping procedures (Figure 1). These arose from needing to instantiate the CSM within a connectionist model. However, this may be also considered an advantage; it forced a novel prediction about the association of CSM beliefs in memory (O'Reilly & Munakata, 2000). Third, the simulation focused on the cognitive representation of illness and omitted or simplified wider aspects of the CSM, including emotional responses, coping representations, and dynamic feed-back from coping evaluation. Emotions may activate beliefs affecting the wider cognitive representation and/or impact coping decisions directly. Dynamic feedback from coping may require a *recurrent* system that aims to predict the utility of specific strategies based on prior experience of the way events string together - generating coping goals (desired outcomes) and assessing the extent to which specific coping procedures achieve expected outcomes. Thus, feedback from coping outcomes may influence treatment adherence, lead people to try different strategies, or even undermine (change) the illness representation because a credible treatment fails to deliver anticipated outcomes (Kucukarslan, 2012). Whilst full feedback is not pursued herein, Simulation 2 did examine how illness schemas evolve over time in response to changing symptom interpretations. Further complexities of coping representation may also need to be looked at, such as differentiating between action planning (anticipation of when, where and how to act) and coping planning (anticipation of obstacles and how to deal with them) (Sniehotta, Schwarzer, Scholz & Schüz, 2005). In sum, addressing complex issues relating to both illness and coping representation provides opportunities for future modeling, perhaps using different network structures or learning rules (for example, recurrent systems for exploring sequences in decision-making or Bayesian networks for modeling representations of cause and effect). Notwithstanding this, we view the current work as an initial step along this road, highlighting the schematic nature of illness representation, the potential for a tipping point in representation change and the generation of counterfactual beliefs. Future work looking at complex associations between illness and coping representation is likely to lead to further conceptual insights and hypotheses.

Finally, one should consider whether connectionist networks are a valid approach for modeling illness beliefs. Our starting point was that beliefs are events in memory which become linked through association. We examined this using a multi-purpose (connectionist) network applying a generalized *delta learning rule*. This models mental association through experience, mirroring the *Rescorla-Wagner* model of reinforcement (O'Reilly & Munakata, 2000; Rescorla & Wagner, 1972). Furthermore, connectionist models of memory have advantages over other approaches: they describe process mechanics (no 'black boxes'), assumptions are explicit, they model self-organization (no *homunculi*), and demonstrate 'brain-like' characteristics such as parallel processing, fault/noise tolerance and content addressability (McLeod, Plunkett & Rolls, 1998; O'Reilly & Munakata, 2000). As such, we consider connectionist networks as a useful framework for exploring low-level processing of illness representation. A network of the sort used herein can be readily modeled in off-the-shelf software (e.g. MATLAB). Whilst modeling requires familiarization with the concepts and program to run, we hope our work encourages other health psychologists to try their own explorations. Future refinements may apply other learning protocols for exploring process complexities such as the influence of feedback on prototype development.

In conclusion, the connectionist network successfully modeled the generation of CSM illness representations. Our analysis suggests people automatically construe representations, in the moment, from knowledge stored in memory, with well-learned information being highly accessible. Illness representations comprise sets of related beliefs organized as schemas. Their schematic nature may mean situations are wrongly interpreted due to incorporation of inappropriate component beliefs. Common questionnaire approaches are possibly inadequate for detecting automaticity in illness representation; alternative assessment tools may be required. The need for further conceptual exploration of complexities surrounding coping and dynamic feedback is highlighted; the current work provides a foundation for this.

References

Bauman, L., Cameron, L.D., Zimmerman, R., & Leventhal, H.. (1989). Illness representations and matching labels with symptoms. *Health Psychology*, 8, 449-469.

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 253-285). New York: Cambridge University Press.

Bean, D., Cundy, T. & Petrie, K. J. (2007). Ethnic differences in illness perceptions, self-efficacy and diabetes self-care. *Psychology and Health, 22*, 787-811. doi: 10.1080/14768320600976240

Broadbent, E., Petrie, K. J., Main, L., & Weinman, J. (2006). The brief illness perception questionnaire. *Journal of Psychosomatic Research, 60*, 631–637. doi:10.1016/j.jpsychores.2005.10.020

Callan, R. (1999). *The essence of neural networks*. Harlow, UK: Pearson.

Clatworthy, J., Hankins, M., Buick, D., Weinman, J., & Horne, R. (2007). Cluster analysis in illness perception research: A Monte Carlo study to identify the most appropriate method. *Psychology and Health, 22*, 123-142. doi:10.1080/14768320600774496

Eiser, J. R. (1994). *Attitudes, chaos and the connectionist mind*. Oxford, UK: Blackwell.

Eiser, J. R., Fazio, R. H., Stafford, T., & Prescott, T. J. (2003). Connectionist simulation of attitude learning: Asymmetries in the acquisition of positive and negative evaluations. *Personality and Social Psychology Bulletin, 29*, 1221–1235. doi: 10.1177/0146167203254605

Eiser, J. R., Stafford, T. & Fazio, R. (2009). Prejudiced learning: A connectionist account. *British Journal of Psychology, 100*, 399-413. doi: 10.1348/000712608X357849

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion, 15*, 115-141. doi: 10.1080/0269993004200024

Fiske, S. T. & Taylor, S. E. (2013). *Social cognition: From brains to culture (2nd Edition)*. London: Sage.

French, D.P., Senior, V., Weinman, J., & Marteau, T.M. (2001). Causal attributions for heart disease: A systematic review. *Psychology & Health, 16*, 77-98. doi: 10.1080/08870440108405491

Hagger, M.S., & Orbell, S. (2003). A meta-analytic review of the common-sense model of illness representations. *Psychology and Health, 18*, 141-184. doi: 10.1080/088704403100081321

Henderson, C. J., Hagger, M. S. & Orbell, S. (2007). Does priming a specific illness schema result in an attentional information-processing bias for specific illnesses? *Health Psychology, 26*, 165-173. doi: 10.1037/0278-6133.26.2.165

Henderson, C. J., Orbell, S. & Hagger, M. S. (2009). Illness schema activation and attentional bias to coping procedures. *Health Psychology, 28*, 101-107. doi: 10.1037/a0013690

Horowitz, C. R., Rein, S. B. & Leventhal, H. (2004). A story of maladies, misconceptions and mishaps: Effective management of heart failure. *Social Science & Medicine, 58*, 631-643. doi:10.1016/S0277-9536(03)00232-6

Kucukarslan, S.N. (2012). A review of published studies of patients' illness perceptions and medication adherence: Lessons learned and future directions. *Research in Social and Administrative Pharmacy, 8*, 371-382. doi: 10.1016/j.sapharm.2011.09.002

Kwate, N.O.A., Thompson, H.S., Valdimarsdottir, H.B., & Bovbjerg, D.H. (2005). Brief report: Etiological attributions for breast cancer among healthy African American and European American women. *Psycho-Oncology, 14*, 421-425. doi: 10.1002/pon.905

Lau, R.R., Bernard, T.M., & Hartman, K.A. (1989). Further explorations of common-sense representations of common illnesses. *Health Psychology, 8*, 195-219.

Leventhal, H., Diefenbach, M., & Leventhal, E.A. (1992). Illness cognition: Using common sense to understand treatment adherence and affect cognition interactions. *Cognitive Therapy and Research, 16*, 143-163.

Leventhal, H., Leventhal, E. A. & Breland, J. Y. (2011). Cognitive science speaks to the "common-sense" of chronic illness management. *Annals of Behavioral Medicine, 41*, 152-163. doi: 10.1007/s12160-010-9246-9

Leventhal, H., Leventhal, E. A. & Contrada, R. J. (1998). Self-regulation, health, and behavior: A perceptual-cognitive approach. *Psychology and Health, 13*, 717-733. doi: 10.1080/08870449808407425

Leventhal, H., Meyer, D., & Nerenz, D. (1980). The common sense representation of illness danger. In S. Rachman (Ed.), *Medical psychology* (pp. 7-30). New York: Pergamon Press.

Leventhal, H., Nerenz, D. R., & Steele, D. J. (1984). Illness representations and coping with health threats. In A. Baum, S. E. Taylor, & J. E. Singer (Eds.), *Handbook of psychology and health: Vol. IV* (pp. 219–252). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lowe, R., Porter, A., Snooks, H., Button L. & Evans, B. A. (2011). The association between illness representation profiles and use of unscheduled urgent and emergency health-care services. *British Journal of Health Psychology, 16*, 862-879. doi: 10.1111/j.2044-8287.2011.02023.x

MATLAB 2013b [Computer software]. The MathWorks, Inc., Natick, MA.

McAndrew, L. M., Musumeci-Szabó, T. J., Mora, P. A., Vileikyte, L., Burns, E., Halm, E. A., Leventhal, E. A. & Leventhal, H. (2008). Using the common sense model to design interventions for the prevention and management of chronic illness threats: From description to process. *British Journal of Health Psychology, 13*, 195-204. doi: 10.1348/135910708X295604

McLeod, P., Plunkett, K. & Rolls, E. T. (1998). *Introduction to connectionist modelling of cognitive processes*. Oxford, UK: Oxford University Press.

McSharry, J., Moss-Morris, R., & Kendrick, T (2011). Illness perceptions and glycaemic control in diabetes: A systematic review with meta-analysis. *Diabetic Medicine, 28*, 1300-1310. doi: 10.1111/j.1464-5491.2011.03298.x

Meyer, D., Leventhal, H., & Gutmann, M. (1985). Common-sense models of illness: The example of hypertension. *Health Psychology, 4*, 115-135.

Moss-Morris, R., Weinman, J., Petrie, K.J., Horne, R., Cameron, L., & Buick, D. (2002). The revised illness perception questionnaire (IPQ-R). *Psychology and Health, 17*, 1-16.

O'Reilly, R. C. & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. London: The MIT Press.

O'Reilly, R. C., Munakata, Y., Frank, M. J., Hazy, T. E., & Contributors. (2012). *Computational Cognitive Neuroscience*. Retrieved from https://grey.colorado.edu/CompCogNeuro/index.php/CCNBook/Book

Rescorla, R. A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton Century Crofts.

Rumelhart, D.E., Hinton, G. E. & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations* (pp. 318-362). London: MIT Press.

Shoda, Y. & Mischel, W. (1998). Personality as a stable cognitive-affective activation network: Characteristic patterns of behavior variation emerge from a stable personality structure. In S. J. Read & L. C. Miller (Eds.), *Connectionist models of social reasoning and social behavior* (pp. 175-208). London: Lawrence Erlbaum Associates.

Smith, E. R. (1996). What do connectionism and social psychology offer each other? Journal of *Personality and Social Psychology, 70*, 893–912.

Smith, E. R. (2009). Distributed connectionist models in social psychology. *Social and Personality Psychology Compass 3*, 64–76. doi: 10.1111/j.1751-9004.2008.00160.x.

Sniehotta, F. F., Schwarzer, R., Scholz, U. & Schüz, B. (2005). Action planning and coping planning for long-term lifestyle change: Theory and assessment. *European Journal of Social Psychology, 35*, 565-576. doi: 10.1002/ejsp.258

Van Overwalle, F. (1998). Causal explanation as constraint satisfaction: A critique and a feedforward connectionist alternative. *Journal of Personality and Social Psychology, 74*, 312-328. doi: 10.1037/0022-3514.74.2.312.

Van Overwalle, F., & Siebler, F. (2005). A connectionist model of attitude formation and change. *Personality and Social Psychology Review, 9*, 231–274. doi: 10.1207/s15327957pspr0903_3.

Van Overwalle, F. & Van Rooy, D. (1998). A connectionist approach to causal attribution. In S. J. Read & L. C. Miller (Eds.), *Connectionist models of social reasoning and social behavior* (pp. 143-171). London: Lawrence Erlbaum Associates.

Weinman, J., Petrie, K.J., Moss-Morris, R., & Horne, R. (1996). The illness perception questionnaire: A new method for assessing the cognitive representation of illness. *Psychology and Health*, *11*, 431-445. doi: 10.1080/08870449608400270

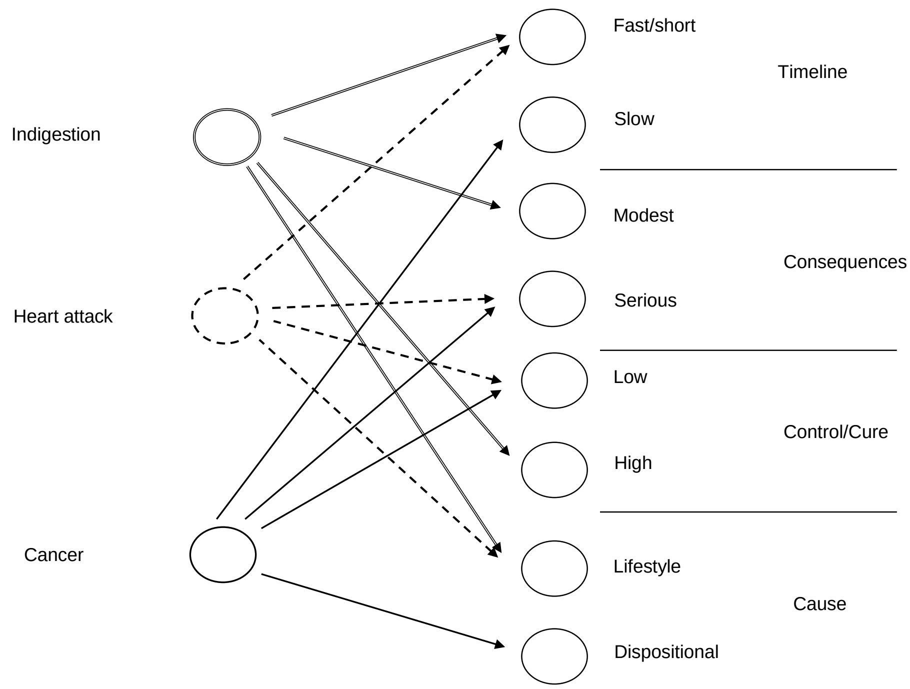

Figure 1a. Association between illness labels and illness beliefs.
Note: Lines show trained associations; for clarity, line styles vary according to illness label.

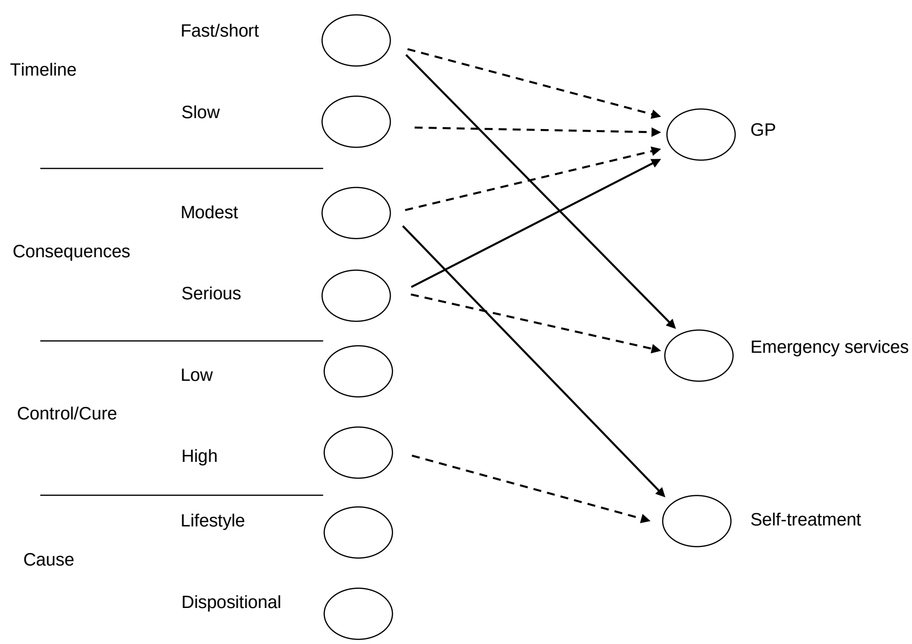

Figure 1b. Associations between illness beliefs and coping procedures.
Note: Lines show trained associations. Line styles vary according to the level of association with solid lines denoting stronger associations than dotted lines.

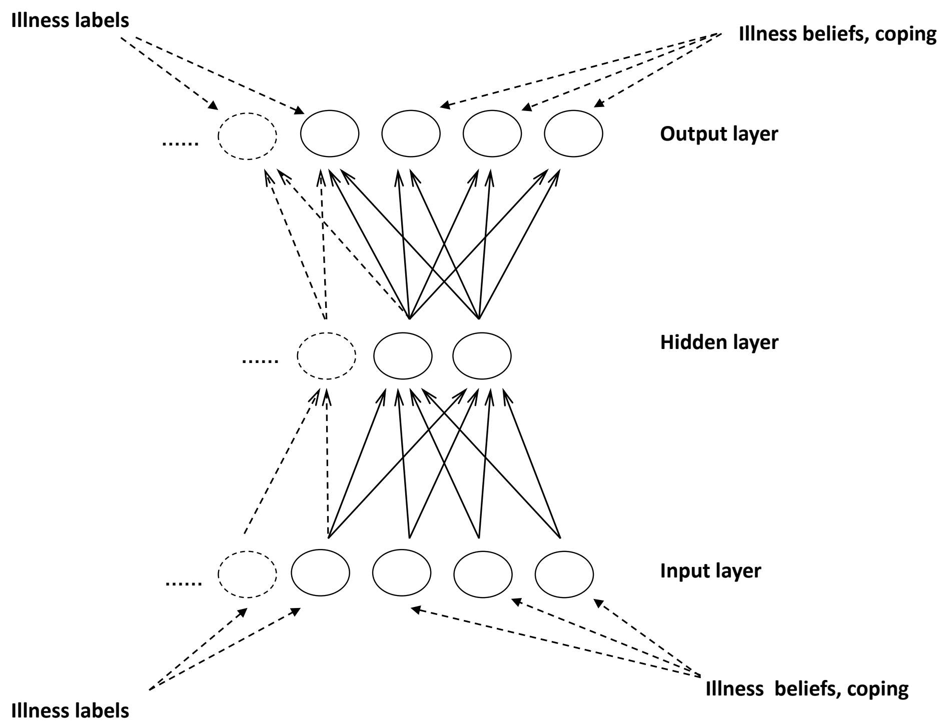

Figure 2. Structure of the autoencoder network
Note: Circles represent nodes arranged into layers, and arrows represent weighted connections. Due to space limitations, the complete network cannot be shown; weights and connections in dotted lines symbolize those omitted. The very bottom and top labels serve to indicate that representations on the output nodes aim to mirror representations on the input nodes.

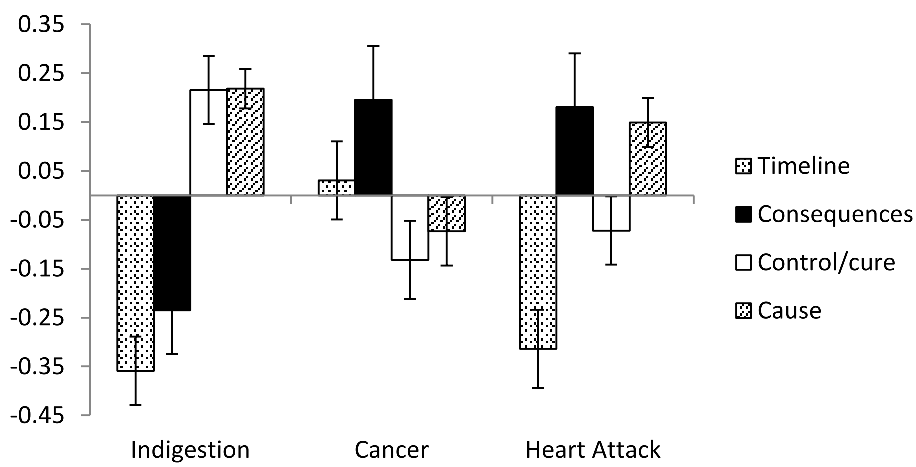

Figure 3. Illness representation profile generated by the connectionist network when presented with a given illness label.
Note: Positive (negative) values represent: long (short) timeline, severe (modest) consequences, high (low) control/cure, lifestyle/behavioral (internal/dispositional) cause

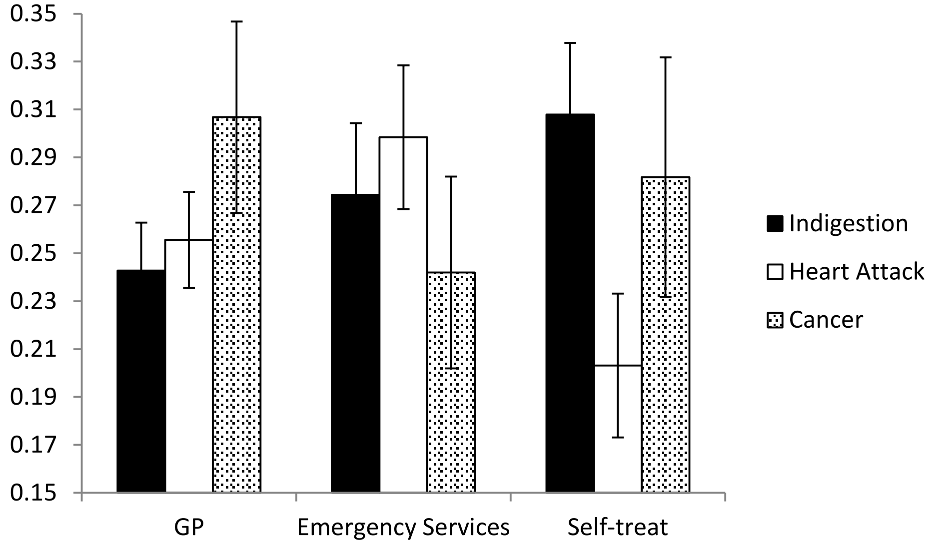

Figure 4. Coping procedures activated by the connectionist network when presented a given illness label

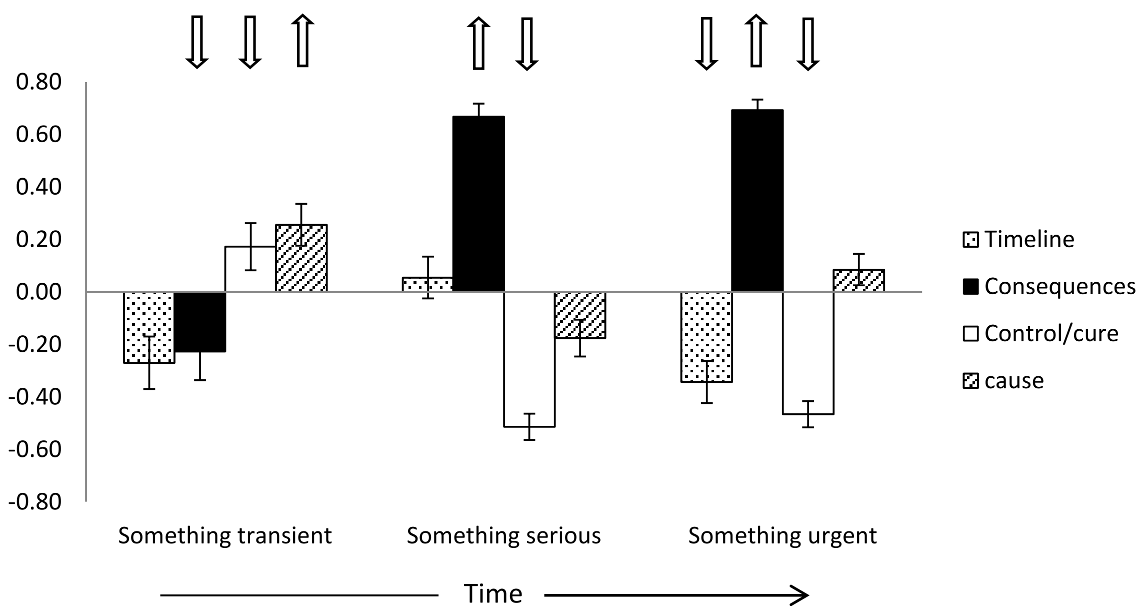

Figure 5. Illness representation profile generated during an unfolding illness situation. The network was presented with select illness beliefs but no illness label.
Note: Columns represent the network's activation at output with positive (negative) values represent: long (short) timeline, severe (modest) consequences, high (low) control/cure, lifestyle/behavioral (internal/dispositional) cause. Arrows denote beliefs activated at input with upward (downward) arrows indicating positive (negative) activation of beliefs at input.

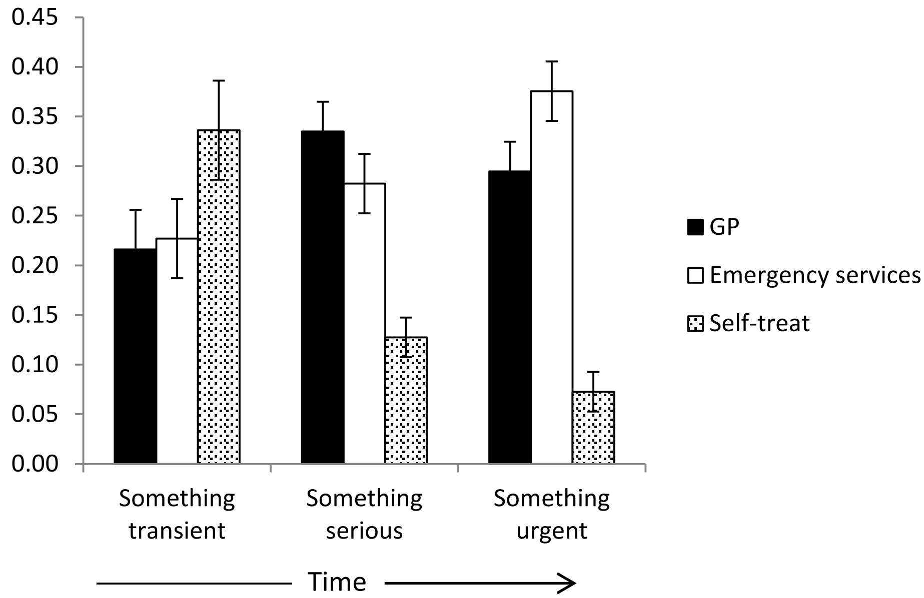

Figure 6. Coping procedures activated for each phase of the illness as it unfolds over time

Table 1
Proportion of pairings for training associations between illness beliefs and a) specific illness labels, and b) specific coping procedures

	Illness Beliefs							
	Timeline		Consequences		Control/Cure		Cause	
	Fast	Slow	Modest	Serious	Low	High	Lifestyle	Disposition
Illness-labels-to-beliefs								
Indigestion	3	0	3	0	0	3	3	0
Heart Attack	3	0	0	3	3	0	3	0
Cancer	0	3	0	3	3	0	0	3
Illness-beliefs-to-coping								
Consult GP	3	3	3	5	0	0	0	0
Emergency Services	5	0	0	3	0	0	0	0
Self-treat	0	0	5	0	0	3	0	0