Cognitive Modeling for Computational Epidemiology

Peter Pirolli[0000-0002-9018-4880]1, Archna Bhatia[2[0000-0002-0593-8212]1], Konstantinos Mitsopoulos[0000-0002-0076-2334]2, Christian Lebiere[0000-0003-4865-3062]2, Mark Orr[0000-0003-2684-4722]3

1 Institute for Human and Machine Cognition, 40 South Alcaniz St., Pensacola, FL 32502, USA
2 Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA
3 University of Virginia, Charlottesville, VA 22904, USA

lncs@springer.com

Abstract. Strategic response options to the COVID-19 pandemic have been greatly influenced by predictive epidemiological models. Effects of non-pharmaceutical interventions (NPIs; such as mask wearing) unfortunately are based on an abundance of very large uncertainties around the extent to which the population adopts risk reducing behaviors. The effects of NPIs appear to have large heterogeneity across regions, subgroups, and individual mindsets and capabilities. We hypothesize that these uncertainties can be improved with higher-fidelity computational modeling of the social-psychological reactions of individuals, groups, and populations. We build up the ACT-R theory and Instance-Based Learning Theory to formulate psychologically valid agents and develop a framework that integrates multi-level cognitive and social simulation with information networks analysis, and epidemiological predictions. We present initial results from analyses of beliefs and sentiments about COVID-19 NPIs induced from online social media that can provide inputs to seed and validate cognitive agents. We present illustrations of cognitive model hypotheses about the dynamics of behavior change in response to intentions, attitudes, messaging, and source credibility. We present an example of social networks propagating attitude change in response to NPIs.

Keywords: COVID-19, ACT-R, epidemiological models.

1 Introduction

Strategic response options to the COVID-19 pandemic have been greatly influenced in the U.S., and elsewhere by predictive epidemiological models [1], which sometimes include individual- or agent-based models (ABMs). Response options for non-pharmaceutical interventions (NPIs; such as wearing masks or social distancing) unfortunately are based on an abundance of very large uncertainties around the likely effectiveness of different policies on risk reducing behaviors. The effects of NPIs (e.g., masks, social distancing) on important factors such as $R(t)$ (reproduction rate), are typically modeled using imprecise assumptions that appear to have large heterogeneity across regions,
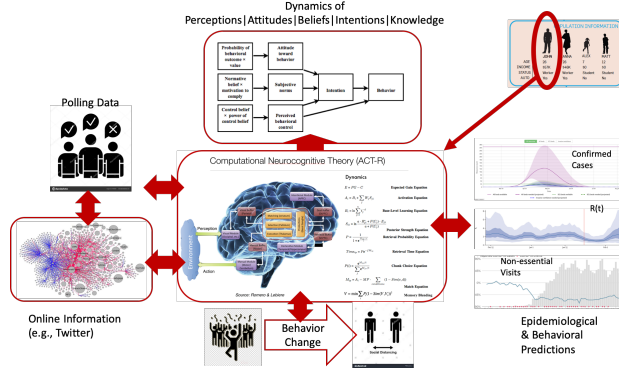
Fig. 1. The conceptual framework: ACT-R agents (center) are used to model behavior change (e.g., social distancing; bottom) in response to information dynamics (left), which have dynamical effects on the perceptions, attitudes, beliefs, intentions, and knowledge (top) influencing decisions and actions (right).

subgroups, and individual mindsets and capabilities, hence producing large imprecision in predicted attack rates and peaks. As the pandemic progresses, decision-makers have needed to move from a blanket approach to NPIs to more precise, sustainable restrictions tailored to geographical regions, social groups, households, and individuals. Consequently, there a need for more *precise* and *accurate* models, and especially models that make predictions about the long-term effects and sustainability of *new kinds of behavior change* (e.g., to wear masks to restaurants or places of worship). To make progress on these challenges, we are researching novel computational models of human responses to epidemics and NPIs during this global pandemic. In this paper, we present a framework that integrates multi-level cognitive and social simulation with information networks analysis, and epidemiological predictions. We present initial results from analyses of beliefs and sentiments about COVID-19 NPIs induced from online social media that can provide inputs to seed and validate cognitive agents. We present illustrations of cognitive model hypotheses about the dynamics of behavior change in response to intentions, attitudes, messaging, and source credibility. We present an example of social networks propagating attitude change in response to NPIs.

2 The COVID Simulation Framework

2.1 Extending Computational Approaches to Epidemiological Prediction

Humans change behavior in response to epidemics. The simulation of infectious disease in populations, in general, assumes that behavioral change is driven by information—information about risk coming directly from epidemic markers (e.g., daily case count) or social cues, perceived risk mediated by psychological constructs, imitation of others behavior (not necessarily indicating risk) due to social norms/influence, and feedback from the results of prior behavior (e.g., an individual's efficacy of the flu shot in prior years). Such assumptions about behavior change are incorporated into both population-level and individual-level modeling approaches (agent-based modeling as the latter). The implementation of behavior-change in simulation models of infectious disease ranges from the use of switches (e.g., change behavioral state given a set of conditions) to more graded behavioral responses (e.g., a linear function between information and behavior change) and sometimes incorporated changes in

network contact structures. Further, some functional specifications between information and behavioral response are quite sophisticated, e.g., game-theoretic models that incorporate social learning. However, to date, we do not know of any infectious disease simulations that incorporate sophisticated computational models of behavioral response that are grounded in psychological first principles or cognitive science; i.e., the union with PVAs has yet to be realized.

We have introduced a general approach for Reciprocal Constraints Paradigm (RCP) for simulating social and individual cognitive systems [14]. Central to this ABM approach (Figure 1) is the use of the ACT-R cognitive architecture [4, 5] or other cognitive architectures (e.g. ,Soar) or approaches from computational psychology (ref Read's work). to model individual agents immersed in social context and interaction (e.g., contact networks) and information consumption and production (e.g., social media, mass media, government communications). This will introduce an unprecedented degree of psychological fidelity to epidemiological modeling.

2.2   The Potential of Psychologically Valid Agents

Our approach (Figure 1) focuses on Psychologically Valid Agents (PVAs) that can simulate individuals with a range of attitudes, beliefs, and credibility assessments that determine their intentions and decisions in response to NPIs (e.g., to shelter-in-place vs go to the beach). Populations of PVAs, in an ABM framework, will be generated using novel techniques that exploit online media and COVID-19 datasets (including polling and epidemiological data). Models can be tested in multiple ways. For instance, we can use the PVAs to generate and test predictions about differences in epidemiological outcomes arising from the natural experiments across US regions that differ in public health interventions and public mindsets. This can be achieved by embedding PVAs in existing ABM epidemiological models. We can also test forecasts about observable behavior that mediates or moderates viral transmission, such as regional mobility and non-essential visit patterns or continuous polls of mask wearing behavior. As of this writing, the framework presented in Figure 1 is notional. Here, we present models and analyses that illustrate key phenomena we expect to model and validate in the future.

2.3   Cognitive Agents to Model NPI Behavior

The cognitive models of mask wearing behavior discussed here were developed using the ACT-R cognitive architecture [4, 5]. Our models are implemented using the ACT-UP framework [15], which makes it flexible as to whether it represents the decisions of a population, community or individual. In this simple simulation, the single model represents the entire population, with the set of instances representing the overall belief system, the action representing the percentage of population engaging in that behavior (mask wearing, in our case), and the outcomes (as well as the situation) representing the system variables resulting from the behavior (e.g., the population level of infection). An individual model would be substantially similar, but involves individual beliefs, the probability of individual action, and individual outcomes. The cognitive architecture provides a computational implementation of a unified theory of cognition that specifies representations and mechanisms for cognitive functions such as

4

perception and attention, memory, decision-making, and action selection. IBL [9] specifies that decisions are based on memories of communications and experiences.

Here, we build upon several existing ACT-R models to formulate a first approximation model of the dynamics of attitudes and behavior in response to experience and information sources regarding NPIs. These models are generally consistent with traditional psychological theories of decision making [13, 17], behavior change [2], mathematical models of attitude change [10], and credibility judgments [6]. There are many advantages that result from using ACT-R to provide a computational formulation of these disparate theories. Amongst these advantages are the integration of multiple factors into a single predictive theory, the modeling dynamics of attitudes and behavior in response to the "dosing" effects of NPI messages, and a foundation for modeling effects of (in)coherence of messaging and sources on credibility.

In our approach, *goal intentions* are goal-like representations that are stored in declarative memory as a kind of prospective memory to be turned into an active goal, in the goal buffer, in response to the right context. Goal intentions vary in magnitude and their retrieval at the right time is dependent on declarative memory activation. Together these affect the likelihood of deciding to pursue the goal (e.g., wear a mask). Having recalled and chosen to pursue an intention to do a behavior (e.g., maintain social distance) one may still require ancillary knowledge and effort to actually carry it out (e.g., skill at estimating a six-foot distance). The decision to pursue one intention versus others that apply in a given situation is assumed to be modeled by the decision-making processes of IBL in ACT-R [9].

Attitudes towards a behavior are assumed to be beliefs and expectancies regarding the behavior's consequences [2] represented as chunks in memory. Attitudes influence intentions [2]. Mathematical models of attitudes [10] and the Theory of Planned Behavior [2, 3] assume that the valence of attitudes is proportional to weighted expectancy-value, i.e., $Attitude \propto \sum belief \cdot evaluation$.

"*Messages*" are the generic term we use to capture any communication (tweet, cable news, website, pamphlet, poster, signage, etc) from an external source (person, media, government official, etc.) [10]. We restrict our interest to those that play a role in stimulating attitude change or reminding/reinforcing intentions. Messages are assumed to have the potential to modulate the valence and magnitude of attitudes.

Source credibility is a cognitive judgement (i.e., subjective) about the source of messages that can modulate message effects on attitudes. Credibility judgments can be conceived as a combination of expertise judgments and trust judgments about the source [16]. In previous work we showed how topic model techniques plus online social network analysis could be used to predict credibility assessments of online twitters users [8], and we propose that such models can be generalized and extended. The functional form of the effect of credibility on attitudes is assumed by some [6, 10] to be multiplicative, $\Delta attitude = source\ credibility \cdot message\ strength$.

3 Analyses and Simulations

3.1 A Pipeline from Data to Cognitive Agent

Since we want to simulate individuals with a range of attitudes, beliefs, and credibility assessments to determine their intentions and decisions in response to NPIs, it is important to acquire information about individuals' attitudes and beliefs from the real world. One of the very commonly used methods for this goal is surveys conducted with individuals belonging to groups of interest as participants. However, conducting surveys is expensive and time consuming, and also, to some degree out of context (not naturalistic in the sense that survey items probe individuals about current and future contexts). It is appealing to use additional methods to gather relevant data. Since user-generated text is already available in the public domain where users express their opinions and preferences, such as in blogs, articles, tweets, reddit posts etc, we hypothesize that it is possible to identify individuals' attitudes and beliefs from the text they generate. Hence, we developed a pipeline to find textual data relevant to the topics (related to attitudes/beliefs) we are interested in. These textual data are then used to extract embedded attitudes and beliefs of the individuals towards various NPIs.

Table 1. Example entity-based analysis

Text in the identified tweet (extended tweet with all the parts):			
Stay home if you feel unwell. If you have a fever, cough and difficulty breathing, seek medical attention. Calling in advance allows your healthcare provider to quickly direct you to the right health facility. This protects you, and prevents the spread of viruses and other infections. Masks can help prevent the spread of the virus from the person wearing the mask to others. Masks alone do not protect against COVID-19, and should be combined with physical distancing and hand hygiene... Follow the advice provided by your local health authority.			
Entities of interest: "mask(s)", "hygiene" **Identified dependency relations for entities of interest (below)** (*Note*: "DR" = Dependency relation; "Direction" = direction of relation)			
1. Entity: Masks Indicator: help DR: nsubj Direction: Dependent	3. Entity: mask Indicator: wearing DR: dobj Direction: Dependent	5. Entity: Masks Indicator: combined DR: nsubjpass Direction: Dependent	7. Entity: hygiene Indicator: combined DR: nmod:with Direction: Dependent
2. Entity: mask Indicator: the DR: det Direction: Governor	4. Entity: Masks Indicator: protect DR: nsubj Direction: Dependent	6. Entity: hygiene Indicator: hand DR: compound Direction: Governor	8. Entity: hygiene Indicator: distancing DR: conj:and Direction: Dependent

Similarly, we searched Twitter data to find tweets relevant to the topics (beliefs) we are interested in. Textual data thus obtained (e.g., tweets or blog posts etc) is parsed. Syntactic relationships are identified between a number of relevant search terms, such as "(face) masks" and "(social) distancing", and other items in the sentences. These items are considered as potential indicators for sentiment of the user towards the NPIs as indicated by the search terms. Entity based sentiment is used to identify the polarity of the sentiment (e.g., positive, neutral, negative) as well as intensity of the sentiment (strong vs weak) toward the entity. The syntactic relationships as well as sentiments can help us determine the attitudes and beliefs of the users towards various NPIs. An example of the kind of analysis involved is shown in Table 1.

For our preliminary version, the potential indicators identified as above are searched in the sentiment lexicons and their sentiment value is assigned to the entity of interest that these potential indicators are in syntactic relationship with to indicate the sentiment of the user towards these entities of interest. The beliefs of the user regarding the entities of interest can be identified using the indicator and its dependency relation with the entity of interest. For example, in chunks 1 and 4 above, a positive sentiment will be assigned to "masks". These chunks indicate that the user holds the following belief regarding masks: "masks help and protect." Detecting such beliefs from user-generated texts provide us with information about the beliefs that individuals hold in the real world. This information can be useful for developing cognitive models of individuals.
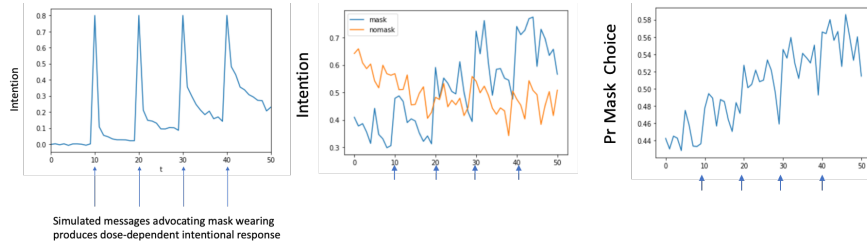

Fig. 2. Left: Simulated effects on intention to wear a mask of hypothetical messages advocating such behavior delivered at given discrete points in time. Center: Competing intentions to wear a mask or not in response to those messages. Right: Probability of choice to pursue mask wearing given competing intentions and NPI messages.

3.2   Attitudes, Intentions, and Behavior in Response to NPI Message "Dosing"

We illustrate ACT-R predictions concerning the chain of effects of NPI messaging→attitude change→intentions→behavior. Declarative memory is populated with experiential and belief chunks that capture outcome expectations and utility of past and potential decisions and behaviors. These expectancy-value attitudes influence the intention to perform a behavior. Figure 2 presents a simulation of the hypothetical effects of receiving messages at discrete points in time concerning the expected value of "mask wearing" with the aim of promoting an intention to "wear a mask." Figure 2 illustrates the ACT-R temporal dose-response of intentions to change behavior in response to messages. The dose-response curve reflects underlying mechanisms of declarative memory: There is a cumulative effect of multiple intention-change messages, but the effects of each specific message decays with time. Figure 2 also presents a simulation in which memory populated with attitudes and intentions to 'not wear a mask' and to 'wear a mask' and messages promoting the latter attitude are received. Messaging drives attitudes and intentions to 'wear a mask' to dominate competing intentions and this has an effect on the decision to pursue the behavior.

Credibility is often defined as a judgment that incorporates assessments of the expertise of, as well as trust in, the source. Here we focus on the role of expertise assessment in credibility. Liao et al. [12] used a Hierarchical Bayesian Model (Latent Dirichlet Allocation; LDA) and a simple online social network analysis to seed ACT-

R declarative memory representations with a set of features that correspond to the latent topics, whose feature values correspond to a probability distribution over those topics. LDA was used to induce a generative probabilistic model for the collection of tweets produced by Twitter users in which each user is represented as a mixture of latent semantic topics, and each topic generates a mixture of words going into tweets [7, 8]. Given a reference domain of expertise (e.g., 'health expert') the topic mixture of an information source (e.g., a twitter user, government spokesperson, etc) can be compared to the topic mixture of the "expert" reference. Canini et al. [8] used a probability-based ranking algorithm based on LDA to predict credibility judgments, whereas Liao et al. relied on ACT-R blending to perform the comparison of an information source to the reference mixture for domain expertise.

We set up ACT-R to have credibility modulate the impact of an NPI messages on attitudes. Figure A illustrates this with a sample of repeated blended retrievals of intentional magnitude in response to messages that, to varying degrees, promote (Valence = +1) or disapprove (Valence = -1) of a specific behavior attitude, as modulated by source credibility. Message sources of low credibility have no effect, whereas high credibility sources have maximal impact.
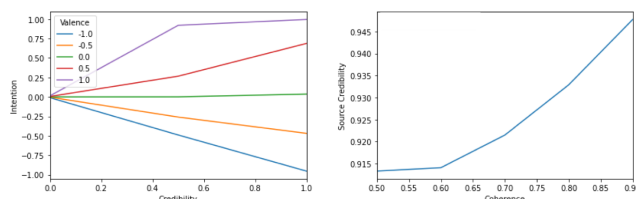

Fig. 3. Left: Modulating effects of source credibility on messages advocating (positive valence) or opposing (negative valence) Right: The credibility of a source having varying degrees of coherence in messaging.

A phenomenon of interest is the effect of the coherence of messages from information sources. We assume that sources that do not provide coherent messages have reduced credibility. Figure B presents a simulation in which a single information source produces messages that primarily generated from one domain of expertise (coherent) or two (incoherent). To do this, we used a single topic mixture (coherent) or two (incoherent) as referent topic mixtures for a source and generated samples of messages that diverged uniformly around those referent mixtures. In Figure B, Coherence = 0.9 means that 9/10 of the source's messages were generated by a single topic mixture (coherent) and Coherence = 0.5 means an even distribution across two topic mixtures (incoherence). The credibility scale is in arbitrary units.

We have not yet addressed the role of trust in credibility. One would expect that sources that have known biases or that have promoted falsehoods would be judged s less credible, and those factors need to be incorporated in future version of our model.

3.3   Modeling the Propagation of Mask-Wearing Behavior

We now show some preliminary modeling efforts on mask wearing. These efforts reflect two parallel efforts that were designed to be integrated. One effort has developed an agent-based framework that captures simple agent-level decisions with respect to mask wearing and the associated infection dynamics. The other effort is an initial ACT-R based SIR-like modeling framework that reflects the relation between

external information, cognitive processing, decision making and the associated infection dynamics. We will review preliminary results from each in turn.

Agent-based Effort

In order to demonstrate the effect of belief change in mask wearing on the spread of a virus, we implemented SIR (susceptible-infectious-recovered) dynamics on a stochastic network that mimics stylized interactions between individuals in a society. Such an approach allows each individual to have a different belief on mask wearing varying over time.
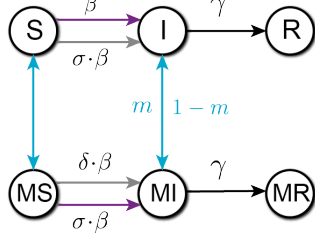

Fig. 4. SIR dynamics with masked and unmasked populations. Nodes can exhibit spontaneous transitions which represent their conversion to a masked or unmasked version of their current state. The rate of this conversion is defined by the current value of their belief on mask wearing. The grey arrows indicate that the node came in contact with a masked node before the transition (e.g. unmasked susceptible comes into contact with an unmasked infected and is converted to an infected node). The purple arrows indicate that the node came into contact with an unmasked node. We assume that a contact between a masked node and unmasked node will discount the transmission rate β by σ. A contact between two masked nodes will discount the transmission even further by δ=σ/2.

We employed a network of 10K nodes and assumed an initial infected population of 1%. Each node can represent an individual or a larger group of individuals that share similar characteristics. The state transition diagram is presented in Fig. 4. An attribute $m \in [0, 1]$ is initially assigned to all nodes indicating the probability of that node wearing a mask. Therefore, the value *m* specifies the rate at which an unmasked node can be converted to a masked one and 1-*m* the rate at which a masked node can be converted to an unmasked one.

For demonstration purposes, we used a monotonically increasing belief update. Thus, at the end of the simulation we expect most nodes to favor mask wearing. Assuming that a node *i* has belief m_i with N immediate neighbors, the belief update is

$$m_i \leftarrow m_i + \frac{\alpha}{N} \sum_{j=1}^{N} m_{j \neq i}$$

where α is a discount factor. The update equation can be substituted by more sophisticated functions that can reflect more complex dynamics that takes into account various cognitive aspects of decision making.

All SIR models were built using the EoN modelling [11]. Fig. 5 show some preliminary results. Panel A shows the dynamics of the simulation when mask wearing is not used because pro-mask beliefs are not distributed in the population. In Panel B, with identical initial conditions, we can see the course of the epidemic when pro-mask beliefs become well distributed over time in the population. This kind of approach, importantly, can serve as a platform for incorporation of cognitive architectures for belief-based decision making.
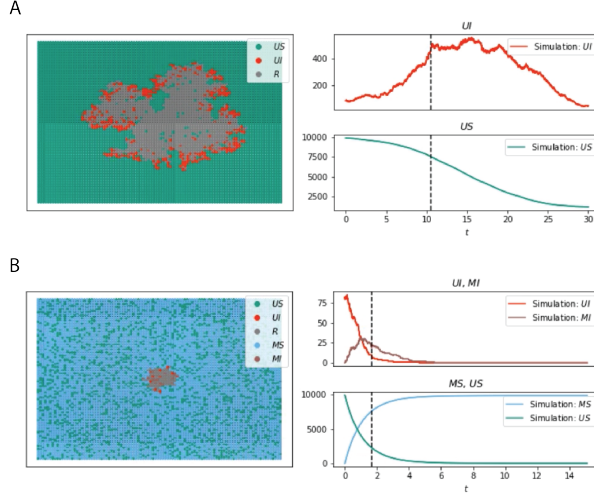
A

B

Fig. 5. Dynamics of mask-wearing propagation.

ACT-R Effort

Figure 6 shows three conditions that instantiate a homogeneous population of 100K agents and differ with respect to the degree to which mask-wearing behavior affects the epidemic dynamics. *Hope* agents discount the effectiveness of mask wearing while *fear* agents are more sensitive to it (neutral is about 50/50). This is a prototype for understanding how person beliefs might interact with the information environment to impact decision-making in epidemics.
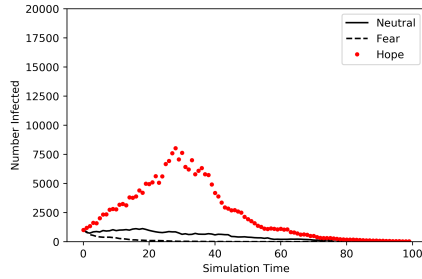

Fig. 6. ACT-R Epidemic dynamics.

4 General Discussion

The dynamics related to NPI compliance and viral transmission arise, in part, from the behaviors of individuals [18]. We argue that we need computational simulation models that generate macro phenomena from the micro dynamics with fidelity to both psychological processes and social structure. For decision-makers, a useful model of the efficacy of implementation of NPIs (such as mask wearing) will depend crucially on locale, pre-existing social norms and networks, individual attitudes, intentions, perceived behavior control, knowledge, skills, and influence from the information ecology. For instance, lock-down policies aimed at reducing elderly infections appear to vary geographically in effectiveness depending on proportion of multi-generational households, preexisting credibility judgments about expertise and media, online social network embedding, and the knowledge and experience needed to comply (e.g., proper hand washing, handling of delivered packages, social distancing). We build upon the Reciprocal Constraints Paradigm for simulating social and individual cognitive systems and embed these within epidemiological ABMs. Psychological Valid Agents will expand upon prior ACT-R models of decision-making and behavior-change and predict individual-level responses and resulting population dynamics. In future work will use online media and other datasets to seed populations of PVAs that simulate populations of the selected regions, and validate these against compliance behavior and epidemiological data.

10

References

1. Adam, D. Special report: The simulations driving the world's response to COVID-19 *Nature*, 2020.
2. Ajzen, I. The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, *50*. 179-211.
3. Ajzen, I. and Kruglanski, A.W. Reasoned action in the service of goal pursuit. *Psychological Review*, *126* (5). 774-786.
4. Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C. and Qin, Y. An integrated theory of mind. *Psychological Review*, *11* (4). 1036-1060.
5. Anderson, J.R. and Lebiere, C. *The atomic components of thought*. Lawrence Erlbaum Associates, Mahwah, NJ, 1998.
6. Birnbaum, M.H. Source credibility in social judgment: Bias, expertise, and the judge's point of view. *Journal of Personality and Social Psychology*, *37* (1). 48-74.
7. Blei, D.M., Ng, A.Y. and Jordan, M.I. Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*. 993-1022.
8. Canini, K., Suh, B. and Pirolli, P., Finding credible information sources in social networks based on content and social structure. in *IEEE International Conference on Social Computing, SocialCom 2011*, (Boston, MA, 2011), IEEE.
9. Gonzalez, C., Lerch, J.F. and Lebiere, C. Instance-based learning in dynamic decision making. *Cognitive Science*, *27*. 591-635.
10. Hunter, J.E., Danes, J.E. and Cohen, S.H. *Mathematical models of attitude change: Volume 1*. Academic Press, Orlando, FL, 1984.
11. Kiss, I.Z., Miller, J.C. and Simon, P.L. *Mathematics of epidemics on networks*. Springer, Cham, 2017.
12. Liao, Q.V., Pirolli, P. and Fu, W. An ACT-R model of credibility judgment of micro-blogging Web Pages. in *Proceedings of the International Conference on Cognitive Modeling (ICCM 2012)*, Universitätsverlag der TU Berlin, Berlin, Germany, 2012, 103-108.
13. Luce, R.D. *Individual choice behavior*. Wiley, New York, 1959.
14. Orr, M.G., Lebiere, C., Stocco, A., Pirolli, P., Pires, B. and Kennedy, W.G. Multi-scale resolution of neural, cognitive and social systems. *Computational and Mathematical Organization Theory*, *25* (1). 4-23.
15. Reitter, D. and Lebiere, C. Accountable modeling in ACT-UP, a sclable, rapid-prototyping ACT-R implementation. in Salvucci, D.D. and Gunzelmann, G. eds. *Proceedings of the 10th International Conference on Cognitive Modeling*, Drexel, Philadelphia, PA, 2010, 199-204.
16. Rieh, S.Y. and Danielson, D.R. Credibility: A multidisciplinary framework. in *Annual Review of Information Science and Technology*, Information Today, Medford, NJ, 2007, 307-364.
17. Slovic, P., Lichtenstein, S. and Fischhoff, B. Decision theory. in Atkinson, R.C., Herrnstein, R.J., Lindzey, G. and Luce, R.D. eds. *Steven's handbook of experimental psychology (Second Edition; Vol. 2)*, Wiley, New York, 1988, 673-738.
18. West, R., Michie, S., Rubin, G.J. and Amlôt, R. Applying principles of behaviour change to reduce SARS-CoV-2 transmission. *Nature Human Behaviour*, *4* (5). 451-459.