Trust and the Dynamics of Testimony*

Wesley H. Holliday

Department of Philosophy, Stanford University
wesholliday@stanford.edu

Abstract

We propose a *dynamic testimonial logic* (DTL) to model communication and belief change among agents with different dispositions to trust each other as information sources. DTL is an extension of the dynamic epistemic logic approach to belief revision (of van Benthem 2007), with the addition of sources and trust. It is also in the spirit of the modal logic approach to trust (of Liau 2003), with the addition of dynamics for belief change. In the multi-agent framework of DTL, we can represent how communication by an information source leads other agents to revise their beliefs about the world, about the source's beliefs, and about the beliefs of other agents in the source's audience. We can also represent how an agent's uncertainty about whether another agent trusts a source can produce, after communication by the source, uncertainty about what the other agent believes, and how an agent can learn whom a source trusts from the source's communication. To capture these phenomena, we introduce a new class of *testimonial* models and model transformations, for which we give a complete axiomatization. Finally, we describe an application of DTL in modeling a special case of the phenomenon of *information cascade* discussed in the economics literature.

*In D. Grossi, L. Kurzen, and F. Velázquez-Quesada, editors, *Logic and Interactive Rationality. Seminar's Yearbook 2009*, pages 147-178. Institute for Logic, Language, and Information, Universiteit van Amsterdam, 2010.

1 Judgment Aggregation to Information Cascades

As it is modeled formally, judgment aggregation is an instantaneous process: given a group of agents, each with an opinion on some proposition, an aggregation function takes their individual opinions and returns a group opinion, all at once (see List and Puppe 2009). Yet in many contexts—from courtrooms to committees—the protocol is to solicit individual opinions sequentially, not simultaneously. For one example in which the temporal dimension matters, consider what Sorensen (1984) calls the *epistemic bandwagon effect*:

> An expert's epistemic preferences can be justifiably influenced by his knowledge of another expert's preferences. Yet this provides the basis for an epistemic bandwagon effect. For the sake of simplicity, suppose there are three highly respectful experts, 1, 2, and 3, who prior to the roll-call vote are respectively in favour, indifferent, and opposed to a proposition. However, they only learn the others' preferences by their votes. If the roll-call vote is taken in order 1, 2, 3, expert 1 votes in favour. Having learned that another expert favours the proposition, the opinion of 2 is swayed and he too votes in favour. Having learned that two experts favour the proposition, 3 reverses his opinion (since he is highly respectful) and the proposition is unanimously favoured. However, if the roll-call vote is taken in order 3, 2, 1, incremental disclosure of preferences and high respect results in the proposition being unanimously opposed.... Disclosure order bias indicates that epistemic respect is trickier than has been supposed. (49-50)

Sorensen's epistemic bandwagon is a special case of what is known in the economics literature as an *information cascade*. An information cascade is a situation in which the preferences, predictions, decisions, etc., of agents are revealed sequentially, and agents acting later in a sequence follow patterns established earlier in the sequence, even when their private information would otherwise suggest deviating from the pattern. Economists have demonstrated information cascades in experiments with human subjects (Anderson and Holt 1997), and they have used the theory of information cascades to analyze phenomena including herd behavior in financial markets, momentum in political campaigns, and fads in medical practice (Bikhchandani et al. 1992). From this perspective, an epistemic bandwagon is an information cascade in which the actions performed sequentially are announcements of agent opinions and the cause of the cascade is epistemic respect among agents.

While Sorensen uses the notion of epistemic respect, a more standard notion is that of *epistemic trust*, understood as trust in another agent's judgment on the truth of a proposition. In general agents learn of each other's judgments via *testimony*, understood in the broad sense of "saying or affirming something in an apparent attempt to convey (correct) information" (Audi 1997, p. 405). A number of authors have stressed the importance in science and mathematics of epistemic trust in the testimony of others, not only between laypeople and experts but also among experts themselves (Hardwig 1985; 1991, Geist et al. 2010). In the multi-agent systems literature, formal models of epistemic trust have been developed using modal logic (Demolombe 2001; 2004, Liau 2003).

In this paper we introduce a *dynamic testimonial logic* (DTL) to model belief change over sequences of testimony among agents with different dispositions to trust each other as information sources. There are several motivations for the framework of DTL. First, DTL extends standard dynamic epistemic logic (DEL) (see van Ditmarsch et al. 2008) by explicitly representing *sources* of information, in such a way that the identity of a source affects how agents revise their beliefs in response to communication from the source. Second, DTL extends the DEL approach to belief revision (of van Benthem 2007) by representing the "missing link" between communication and belief revision, the *acceptance* of information. In DTL, whether an agent accepts information depends on whether the agent trusts the source. Third, while existing logics of epistemic trust are static logics, providing a snapshot view of the information and trust relations in a multi-agent setting, DTL is a dynamic logic, capable of modeling the role of trust in temporal phenomena such as information cascades. Finally, DTL models truly multi-agent belief revision, for it models not only how agents revise their beliefs about other agents' beliefs, but also how agents perform different types of belief revisions in response to the same informational event.

In the rest of Section 1 we provide the conceptual basis of DTL, drawing distinctions between *public announcement* and *testimony* and between the *doxastic* and *testimonial reliability* of agents. In Section 2 we review the logic of belief underlying DTL, the *conditional doxastic logic* of Baltag and Smets (2008). Turning to dynamics in Section 3, we review the approach to iterated belief revision of van Benthem's (2007) *dynamic logics of belief upgrade*. Given these foundations we introduce DTL in Section 4, adding *testimonial records* and *authority relations* to our models to capture agents' epistemic trust in the testimony of others. We then represent testimony dynamically by transformations on these enriched models. Finally, we describe an application of DTL in modeling epistemic bandwagons in Section 5, and we conclude with directions for further research in Section 6. The Appendix contains a complete axiomatization for DTL.

1.1 Trust and Authority in Testimony

Consider those experts on whose authority you would be willing to believe a proposition φ. We will say that you "trust the judgment" of these experts on φ. Among your trusted experts, some may be more authoritative for you than others. If expert 1 testifies that φ, expert 2 testifies that $\neg\varphi$, and you come to believe φ, then 1 is more authoritative for you than 2. If 2 were more authoritative, then you should have come to believe $\neg\varphi$. And if 1 and 2 were equally authoritative, you should not have changed your mind on φ either way, or you should have suspended judgment on φ altogether. The same points apply if 1 and 2 are groups of experts, rather than individuals.

If 1 is more authoritative for you than 2, then after 1 testifies that φ, you no longer trust 2 on $\neg\varphi$, in the sense that you will no longer believe $\neg\varphi$ on the authority of 2, something you might have done before 1 testified. However, if another expert 3 joins 2 in testifying that $\neg\varphi$, you may believe $\neg\varphi$ on the authority of 2 *together with* 3, though perhaps not on the authority of either of them individually. For our formalization we will assume that each agent has a ranking of the authority of other (sets of) agents, allowing incomparabilities (for how agents might rationally arrive at such rankings, see Goldman 2001).

Given our intuitive picture of trust and authority, our goal is to develop a model of testimony that addresses the following questions about what happens when an agent i testifies that φ. First, what do agents in i's audience come to believe about φ and related propositions? This question subdivides into two others: what determines whether agents in i's audience *accept* φ, and for those who accept φ, how do they revise their beliefs given this new information? Second, what do agents in i's audience come to believe *about i's beliefs*? Finally, what do agents come to believe *about the beliefs of other agents in i's audience*?

1.2 Testimony vs. Public Announcement

To identify the information provided by testimony, it is useful to compare testimony with *public announcement*, the classic case of an informational event in dynamic epistemic logic. For our purposes, the crucial difference between testimony and announcement is that while announcements are typically thought to come from an anonymous external source, testimony will always come from one of the agents within our model.

What difference does the individual source of testimony make? Let us make two assumptions about the kind of testimony in question. First, suppose testimony is *public*: the identity of the testifier and the content of the testimony is information available to all agents. Second, suppose testimony is heard under the *presumption of sincerity*: if an agent i testifies that φ, all other agents come to believe that i believes φ. As a consequence of these assumptions, when an agent i testifies that φ, other agents will acquire the information *that i believes φ*. But then what is the difference between a truthful *public announcement* that i believes φ and i's *own public testimony* that φ, if both provide the information that i believes φ?

One difference is that i's testimony provides *more* information:[1] it provides the information *that i is willing to publicly assert φ*. As we might say, i is willing to "go on the record" for φ. If i is the kind of agent who publicly asserts a proposition only if she has conducted a thorough inquiry into its truth, then the information that i is willing to publicly assert φ is vital information. A truthful public announcement (from no particular agent) that i believes φ does not provide this vital information. For it may be that i believes many propositions, while she only has the time and resources to investigate some few of them in such a way that she would be willing to make public assertions about them.

We can now make a distinction between two ways in which one agent might judge another to be "reliable" on the truth of a proposition. Agent j judges agent i to be *doxastically reliable* on φ just in case if j were to learn that i believes φ (and nothing stronger), then j would believe φ (cf. Demolombe 2004, on "competence"); j judges i to be *testimonially reliable* on φ just in case if j were to learn that i sincerely testified that φ, then j would believe φ. The important point is that judgments of doxastic and testimonial reliability may come apart. Suppose that i has expressed a general lack of understanding of economics. Then j might judge i's doxastic reliability on each economic proposition to be low. But suppose that j knows that i would publicly assert a proposition only if i had conducted a thorough inquiry into its truth. Then j might judge i's testimonial reliability on each economic proposition to be high; if i were to ever make a public assertion about economics, j would take it seriously.

As we have defined it, testimonial reliability concerns the reliability of an agent's *sincere* testimony. Yet an agent whose sincere testimony is highly reliable may often be insincere (cf. Cantwell 1998, p. 195). Perhaps i rarely says what she really believes about economics, due to peer pressure or controversy. Even so, when j is confident of i's sincerity, j might consider i's testimony to be highly reliable. We will see that the design of DTL reflects these distinctions.

[1] At this point we are not considering the difference that public announcement is usually conceived as a source of "hard information" that *eliminates* possibilities, while testimony is better conceived as a source of "soft information" that reorders the relative plausibility of possibilities (see van Benthem 2007).

2 Logic for Conditional Belief

In this section we define the logic of belief underlying DTL, the *conditional doxastic logic* (CDL) of Baltag and Smets (2008). CDL is equivalent to the strongest logic for conditional belief introduced by Board (2004, Sec. 3.3), but we follow the presentation of Baltag and Smets. We focus on the semantics of CDL, referring the reader to the cited sources for axiomatizations.

Definition 2.1. Let At be a set of atomic sentence symbols and Agt a set of agent symbols. The language of CDL is defined by:

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \varphi \mid B_i^{\varphi}\varphi$$

where $p \in \mathsf{At}$, $i \in \mathsf{Agt}$.

We adopt the usual definitions of $\vee$, $\rightarrow$, and $\leftrightarrow$ in terms of $\neg$ and $\wedge$. The intended reading of $B_i^{\varphi}\psi$ is "i believes that ψ conditional on φ." Intuitively, the formula $B_i^{\varphi}\psi$ indicates that if i were to receive the information that φ (and nothing stronger), then i would believe ψ.[2]

For the semantics of CDL, we need some preliminary terminology. Where $\leq$ is a binary relation on a set W, a *comparability class* for $\leq$ is a set $C = \{w \in W \mid w \leq v \text{ or } v \leq w\}$ for some $v \in W$. The relation $\leq$ is a *well-preorder* on W if it is reflexive, transitive and every non-empty subset of W has a $\leq$-minimal element. Finally, $\leq$ is *locally well-preordered* on W if for each comparability class $C \subseteq W$, the restriction of $\leq$ to C is a well-preorder on C.

Definition 2.2. A *multi-agent plausibility model* is a triple $\mathcal{M} = \langle W, \leq, V\rangle$ where W is a non-empty set, $\leq$ is a family of locally well-preordered relations $\leq_i \subseteq W \times W$ for each $i \in \mathsf{Agt}$, and $V : \mathsf{At} \rightarrow \mathcal{P}(W)$.

As in epistemic logic, we think of each $w \in W$ as a possible state of the world (according to some agent), where agents may be uncertain about which is the *actual* state of the world. In CDL agents may also consider some states *more plausible* than others, as represented by the plausibility (pre-)ordering $\leq_i$ for each agent i. Following convention we read $w \leq_i v$ as "agent i considers state w at least as plausible as state v," so the *minimal* states in the ordering $\leq_i$ are the *most* plausible states for i. For comparability of states we write $w \sim_i v := w \leq_i v$ or $v \leq_i w$. Since $\sim_i$ is an equivalence relation it partitions

[2]More accurately, i would believe that ψ *was* the case, before i received the information that φ. This qualification is necessary to make sense of satisfiable formulas such as $B_i^{\varphi}(\neg B_i\varphi \wedge \varphi)$.

W into equivalences classes, which for a given $w \in W$ we denote by $\sim_i(w) = \{v \in W \mid w \sim_i v\}$, called the *information cell* of w for i. Agents only compare states that they consider possible, so we take $\sim_i(w)$ to be the set of states that i considers possible according to her information at w. We read $w \sim_i v$ accordingly as "v is accessible for i from w." As usual, the valuation V sets the atomic facts at every state by mapping each $p \in \mathsf{At}$ to a set of states, (by the truth definition) the set of states satisfying p.

Definition 2.3. Given a model $\mathcal{M} = \langle W, \leq, V\rangle$ and state $w \in W$, we define $\mathcal{M}, w \models \varphi$ (φ is true in $\mathcal{M}$ at w) as follows:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$
- $\mathcal{M}, w \models \neg\varphi$ iff $\mathcal{M}, w \not\models \varphi$
- $\mathcal{M}, w \models \varphi \wedge \psi$ iff $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$
- $\mathcal{M}, w \models B_i^{\varphi}\psi$ iff for all $v \in \min_{\leq_i}(\llbracket\varphi\rrbracket_{\mathcal{M}} \cap \sim_i(w)) : \mathcal{M}, v \models \psi$

where we denote the set of most plausible states in $P \subseteq W$ by $\min_{\leq_i} P = \{v \in P \mid v \leq_i u \text{ for all } u \in P\}$ and the *truth set* of φ by $\llbracket\varphi\rrbracket_{\mathcal{M}} = \{u \in W \mid \mathcal{M}, u \models \varphi\}$. If the intended model is clear from the context, we write $\llbracket\varphi\rrbracket$ instead of $\llbracket\varphi\rrbracket_{\mathcal{M}}$.

If φ is true at a state w, we call w a "φ-state." Simply put, the truth definition for conditional belief says that i believes ψ conditional on φ iff all the most plausible φ-states for i are ψ-states. Given this definition, (unconditional) belief and knowledge are derived operators. Where $\top$ is any tautology and $\bot := \neg\top$, we define $B_i\varphi := B_i^{\top}\varphi$, read "$i$ believes φ," $K_i\varphi := B_i^{\neg\varphi}\bot$, read "$i$ knows φ," and $\hat{K}_i\varphi := \neg K_i\neg\varphi$, read "$i$ considers φ possible."

Multi-agent plausibility models contain information about what agents would believe upon learning various facts. They also contain information about what agents would believe upon learning about *other agents' beliefs*. Consider the formula $B_i^{B_k p}p \wedge B_i^{B_k\neg p}\neg p$ $(*)$, which is true if and only if in all the $B_k p$-states that i considers most plausible, k's belief is true, and likewise for the most plausible $B_k\neg p$-states. Intuitively, $(*)$ expresses that i takes k to be *doxastically reliable* on p, in the sense of Section 1.2. Various judgments of doxastic unreliability can be expressed in a similar way. We can even extend this observation to agents' beliefs about the relative doxastic reliability of other agents. For example, a formula such as $B_i^{B_j p \wedge B_k \neg p}p \wedge B_i^{B_j\neg p \wedge B_k p}\neg p$, which is consistent with $(*)$, expresses i's belief in the superior doxastic reliability of j relative to k on p.

Although multi-agent plausibility models contain information about agents' views of the doxastic reliability of other agents, they do not contain information about agents' views of the *testimonial* reliability of other agents. If we were to assimilate k's testimony that φ to a public announcement of $B_k\varphi$, then these models would be sufficiently rich to determine how agents' beliefs change in response to this "testimony." However, as discussed in Section 1.2, k's testimony that φ is not equivalent to a public announcement of $B_k\varphi$, in terms of the information provided. For this reason we will add additional structure to models for DTL in Section 4.

3 Logics for Iterated Belief Revision

Having defined the underlying logic of belief for DTL, we turn to the dynamics. To model belief revision, we follow the approach of van Benthem's (2007) *dynamic logics of belief upgrade*. These logics provide a formalization not only of "one-shot" belief revision, as CDL does, but also of two well-known *iterated* belief revision policies. Moreover, the same style of analysis can be used to provide complete logics for other iterated belief revision policies. Given the versatility of this approach, one is free to choose one's preferred belief revision policies and use the corresponding logic as a dynamic base for DTL.

Definition 3.1. Let At be a set of atomic sentence symbols and Agt a set of agent symbols. The language of belief upgrade is defined by:

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \varphi \mid B_i^{\varphi}\varphi \mid [\uparrow_i \varphi]\,\varphi$$

where $p \in \mathsf{At}$, $i \in \mathsf{Agt}$.

The intended reading of $[\uparrow_i \varphi]\,\psi$ is "after the revision of i's beliefs with the new information that φ, ψ is the case," or more concisely, "after i upgrades with φ, ψ is the case." While van Benthem uses the symbol $\uparrow$ for conservative upgrade in particular, we use $\uparrow$ for an arbitrary belief upgrade action, and we define three particular upgrade actions below.

Models for the logic of belief upgrade are the same multi-agent plausibility models as before.

Definition 3.2. Given a model $\mathcal{M} = \langle W, \leq, V\rangle$, the model $\mathcal{M} \uparrow_i \theta = \left\langle W, \leq^{\uparrow_i \theta}, V\right\rangle$ is obtained by changing the plausibility ordering $\leq_i$ to $\leq_i^{\uparrow\theta}$ as follows. *Conservative upgrade* (Boutilier 1996): in each information cell for i, the most plausible θ-states become most plausible overall, but otherwise the ordering remains the same. *Restrained upgrade* (Booth and Meyer 2006): in each information cell for i, each set of equi-plausible states is split, such that all θ-states in the set become more plausible than all $\neg\theta$-states in the set, but otherwise the ordering remains the same; then the most plausible θ-states become most plausible overall, but otherwise the ordering remains the same. *Lexicographic upgrade* (Nayak 1994): in each information cell for i, all θ-states become more plausible than all $\neg\theta$-states, but otherwise the ordering remains the same.

Definition 3.3. The truth definition for static formulas is that of CDL. The truth definition for belief upgrade is:

- $\mathcal{M}, w \vDash [\uparrow_i \theta]\, \varphi$ iff $\mathcal{M} \uparrow_i \theta, w \vDash \varphi$

According to the definition, to determine whether $[\uparrow_i \theta]\varphi$ is true at w in the initial model we simply check whether φ is true at w in the updated model.

For each of the three upgrade operations defined, one can give *reduction axioms* that allow the rewriting of any formula with upgrade operators as an equivalent formula in the static language of CDL. Given a complete axiomatization for CDL, these reduction axioms provide a complete axiomatization for the dynamic logics of upgrade. Van Benthem (2007) gives reduction axioms for lexicographic and conservative upgrade.[3] The analogous result for restrained upgrade, which we leave as an exercise for the reader, also holds.

In modeling belief change due to testimony, we wish to model not only how agents form new beliefs, but also how agents *suspend belief*. Suppose that agents j and k are equally authoritative in the eyes of i. If j testifies that φ and then k testifies that $\neg\varphi$, one policy for i would be to believe whoever testified first—in this case, agent j. A more sensible policy in a situation where equally authoritative sources conflict would be to suspend belief about φ. Alternatively, i might not perform any belief revision, ignoring the conflicting testimony of j and k. Yet conflicting testimony from *authoritative* sources does not seem to "cancel out" to provide i with no information. Something informative has occurred—two authoritative sources have testified for φ and $\neg\varphi$ respectively—and i's beliefs should reflect this.

Let us add a *suspension operator* $[\downarrow_i \varphi]$ to the language of the previous section. The intended reading of $[\downarrow_i \varphi]\, \psi$ is "after i suspends belief on φ, ψ is the case."

[3] In the multi-agent case, we must add the reduction axiom $[\uparrow_i \theta]\, B_j^{\varphi} \psi \leftrightarrow B_j^{[\uparrow_i \theta]\varphi}\, [\uparrow_i \theta]\, \psi$ for $j \neq i$ to those given by van Benthem.

Definition 3.4. Given a model $\mathcal{M} = \langle W, \leq, V \rangle$, the model $\mathcal{M} \downarrow_i \theta = \left\langle W, \leq^{\downarrow_i \theta}, V \right\rangle$ is obtained by changing the plausibility ordering $\leq_i$ to $\leq_i^{\downarrow \theta}$ as follows. *Conservative suspension*: in each information cell for i, the most plausible θ states and the most plausible $\neg\theta$-states becomes equally plausible and most plausible overall, but otherwise the old ordering remains the same.

Conservative suspension is the suspension counterpart of conservative upgrade. The pair has several natural properties, one of which is a kind of *confluence*: conservative upgrade with φ followed by conservative suspension with φ produces the same result as conservative upgrade with $\neg\varphi$ followed by conservative suspension with $\neg\varphi$. In our example above, this means that the order in which j and k testify does not make a difference to what i believes after she suspends belief on φ (or $\neg\varphi$), which seems intuitive. Other suspension policies are possible, but we will not consider them here.

Definition 3.5. The truth definition for belief suspension is:

- $\mathcal{M}, w \vDash [\downarrow_i \theta]\varphi$ iff $\mathcal{M} \downarrow_i \theta, w \vDash \varphi$

Reduction axioms for conservative suspension are easily obtained by analogy with those for conservative upgrade. A generalization of conservative suspension is the *multi-upgrade* operation $\downarrow_i \{\varphi_1, ..., \varphi_n\}$, defined as follows.

Definition 3.6. Given a model $\mathcal{M} = \langle W, \leq, V \rangle$, the model $\mathcal{M} \downarrow_i \{\theta_1, ..., \theta_n\} = \left\langle W, \leq^{\downarrow_i \{\theta_1, ..., \theta_n\}}, V \right\rangle$ is obtained by changing the plausibility ordering $\leq_i$ to $\leq_i^{\downarrow \{\theta_1, ..., \theta_n\}}$ as follows: in each information cell for i, the most plausible θ_k-states for each k ($1 \leq k \leq n$) become equally plausible and most plausible overall, but otherwise the ordering remains the same.

The advantage of multi-upgrade over simple (conservative) suspension is that with multi-upgrade an agent can suspend belief on θ while making sure to retain or gain belief in φ. The multi-upgrade $\downarrow_i \{\varphi \wedge \theta, \varphi \wedge \neg\theta\}$ accomplishes the desired effect,[4] for which we will find a use in the next section. Given

[4]Multi-upgrade accomplishes the desired effect provided there are states satisfying $\varphi \wedge \theta$ and states satisfying $\varphi \wedge \neg\theta$ in the information cell of interest. If there are no states satisfying $\varphi \wedge \neg\theta$, for example, then the multi-upgrade $\downarrow_i \{\varphi \wedge \theta, \varphi \wedge \neg\theta\}$ amounts to a conservative upgrade by $\varphi \wedge \theta$. To prevent this, one may wish to define multi-upgrade so that if there are no θ_k-states for some k ($1 \leq k \leq n$) in a given information cell, then the operation $\downarrow_i \{\theta_1, ..., \theta_n\}$ does nothing to that information cell. We will not require this safeguard, however, for we will only use multi-upgrade when it is guaranteed that θ_k-states exists for each k (see the definition of $\updownarrow\updownarrow_j^i \varphi$ in Section 4.4.2).

the reduction axioms for conservative suspension, the reduction axioms for multi-upgrade are a straightforward generalization.

We end our discussion of upgrade and suspension by noting that in the multi-agent setting, belief revision via relation change is a kind of *public* belief revision; when one agent's plausibility relation changes, other agents may "notice" the change. To be precise, the following is a validity:

$$\vDash B_j\hat{K}_i\varphi \leftrightarrow [\uparrow_i \varphi] B_jB_i\varphi$$

We return to the issue of the publicity of belief revision in Section 4.4 below.

4 Dynamic Testimonial Logic

In this section we develop the framework of DTL.

4.1 Language of DTL

Definition 4.1. Let At be a set of atomic sentence symbols and Agt a *finite* set of agent symbols. The language of DTL is defined by:

$$\begin{array}{lll} \varphi_0 & := & p \mid \neg\varphi_0 \mid \varphi_0 \wedge \varphi_0 \\ \varphi & := & \varphi_0 \mid \neg\varphi \mid \varphi \wedge \varphi \mid U\varphi_0 \mid B_i^{\varphi}\varphi \mid \mathtt{rec}_i\varphi_0 \mid S \preceq_i^{\varphi_0} S' \mid [!_i\varphi_0]\varphi \end{array}$$

where $p \in \mathsf{At}$, $i \in \mathsf{Agt}$, and $S, S' \subseteq \mathsf{Agt}$.

The language of DTL includes several new types of formulas. The intended reading of $\mathtt{rec}_i\varphi$ is "i is (most recently) on the record as testifying in favor of φ." The intended reading of $S \preceq_i^{\varphi} S'$ is "S' is as (testimonially) authoritative as S on φ for i." We use the abbreviations $S \prec_i^{\varphi} S' := S \preceq_i^{\varphi} S' \wedge S' \npreceq_i^{\varphi} S$ and $S \approx_i^{\varphi} S' := S \preceq_i^{\varphi} S' \wedge S' \preceq_i^{\varphi} S$. We also allow $\emptyset$ to occur in formulas: we read $\emptyset \prec_i^{\varphi} S$ as "S's testimony on φ is *authoritative* for i," $\emptyset \approx_i^{\varphi} S$ as "S's testimony on φ is *unauthoritative* for i," and $S \prec_i^{\varphi} \emptyset$ as "S's testimony on φ is *anti-authoritative* for i." The reason for this choice of terminology will be clear when we turn to the semantics. Finally, the intended reading of $[!_i\varphi]\psi$ is "after i (publicly) testifies that φ, ψ is the case." For simplicity in this version of DTL we consider only testimony on propositional formulas, so agents do not testify about the beliefs or authority of others. It is for this reason that only φ_0 formulas can appear inside testimony operators and in record and authority formulas. Finally, we have added the universal modality U for the purposes of our axiomatization, so we can express the equivalence of two φ_0 formulas in a model (see Appendix).

4.2 Semantics of DTL

Definition 4.2. A *testimonial model* $\mathcal{M} = \langle W, \leq, V, \mathsf{rec}, \preceq \rangle$ is a multi-agent plausibility model together with $\mathsf{rec} : \mathsf{Agt} \times W \to \mathcal{P}(\mathcal{P}(W))$ and a family $\preceq$ of relations $\preceq^P_{i,w} \subseteq \mathcal{P}(\mathsf{Agt}) \times \mathcal{P}(\mathsf{Agt})$ for each $i \in \mathsf{Agt}$, $w \in W$, and $P \subseteq W$.

The *testimonial record* rec records the set of propositions $\mathsf{rec}_i(w)$ to which i has testified at w, where a proposition is now understood as a *set of states* $P \subseteq W$, not a formula. The intuition is that when i testifies that φ, she is claiming that the actual state is among the φ-states. Hence she goes on the record for $[\![\varphi]\!]$ (at every state w, assuming the testimony is public).[5] The *authority relation* $\preceq^P_{i,w}$ encodes i's view at w of the relative authority of (sets of) agents on the proposition P. Since the authority relations are not necessarily total, we do not assume that all (sets of) agents are comparable in authority for an agent.

Definition 4.3. A testimonial model is *legal* iff for every $i \in \mathsf{Agt}$, $S, S' \subseteq \mathsf{Agt}$, $w, v \in W$, and $P \subseteq W$:

1. $\preceq^P_{i,w}$ is a preorder on P.
2. $S \preceq^P_{i,w} S'$ iff $S \preceq^{W \setminus P}_{i,w} S'$.
3. If $S \preceq^P_{i,w} S'$ $(S \neq S')$, then there is a $v \in \sim_i(w)$ such that $v \in P$ and $\sim_k(v) \cap (W \setminus P) \neq \emptyset$ for all $k \in S$.
4. If $w \sim_i v$, then $\preceq^P_{i,w} = \preceq^P_{i,v}$ and $\mathsf{rec}_i(w) = \mathsf{rec}_i(v)$.

The first condition reflects the assumption that the relation of being *as authoritative as* is reflexive and transitive. The second condition states that authority relations are the same for a proposition and its complement, e.g., i considers j authoritative on whether there will be a recession next year if and only if i

[5]This reflects our semantic perspective, from which we ignore syntactic differences of formulas that pick out the same proposition. We are not taking "on the record" in a literal sense, as a matter of *what the testifier said*. If we wished to keep track of such linguistic matters, we would have opted for a syntactic approach whereby agents go on the record for formulas rather than propositions.

considers j authoritative on whether there will *not* be a recession next year. This condition could be dropped, at the expense of complicating the system.[6]

For the third condition, suppose that P is the truth set of some formula φ. Then the condition implies that if there is a group S' that is at least as authoritative as S on φ for i (at w), then i must consider it possible (at w) for the agents in S to sincerely testify for or against φ and yet be mistaken. That is, there must be an i-accessible state at which φ is true but no one in S knows φ and (by the second condition) an i-accessible state at which φ is false but no one in S knows $\neg\varphi$. For if the only i-accessible states at which φ is true (resp. false) are ones at which someone in S knows φ (resp. $\neg\varphi$), then i considers it impossible for the members of S to all sincerely testify against (resp. for) φ and yet be mistaken. But i must consider this possible if there is some S' that is as authoritative or more authoritative on φ than S. In the special case of $S = \emptyset$, the second condition implies that if there is a set S' of agents that is authoritative on φ for i (at w), then i must consider φ possible (at w).[7]

Finally, the fourth condition states that agents have knowledge of their own authority relations and testimonial records, reflecting the assumption that agents have introspective access to their views of other agents and memory of their own past testimony. Additional conditions on authority relations, which for the sake of generality we will not assume as part of the framework of DTL, may be desirable in certain modeling situations. For example, one might consider conditions on the authority of related sets of agents (cf. Cantwell 1998, p. 194). Possibilities include a uniformity condition of the form $S \preceq^P_{i,w} S' \Rightarrow S \cup S'' \preceq^P_{i,w} S' \cup S''$ and various right and left monotonicity conditions, such

[6]Without the second condition, we would have to give $S \preceq^{\varphi}_{i} S'$ a more complicated reading than "S' is as authoritative as S on φ for i," since both $S \preceq^{\varphi}_{i} S'$ and $S' \preceq^{\neg\varphi}_{i} S$ could be true at the same time. One reason to drop the condition is that the principle that j is authoritative on φ for i if and only if j is authoritative on $\neg\varphi$ for i is subject to counterexamples. For example, since studies have shown that people tend to overestimate their driving ability relative to the average driver, if i knows this fact then i might consider j authoritative on whether j is *not* a very good driver but not consider j authoritative on whether j is a very good driver. However, in the case of *expert* testimony, the symmetric authority of the second condition seems plausible (cf. Liau 2003, on "symmetric trust").

[7]Given the second condition, the third condition also implies that if i "knows" φ, then i does not consider any set of agents authoritative on φ. For if i considers S authoritative on φ, then by the second condition i considers S authoritative on $\neg\varphi$, whence by the third condition there is an i-accessible state satisfying $\neg\varphi$, so i does not know φ. We could avoid this consequence by dropping the second condition, but it would not be worth the loss of simplicity in our system. Since we are not interested in agents who have already made up their minds about propositions, but rather in agents who come to believe or disbelieve propositions on the basis of testimony, it is not necessary to express that an agent may know φ and yet still consider others authoritative on φ.

as $S \preceq^P_{i,w} S \cup S'$, $\emptyset \preceq^P_{i,w} S' \Rightarrow S \preceq^P_{i,w} S \cup S'$, and $S' \preceq^P_{i,w} \emptyset \Rightarrow S \cup S' \preceq^P_{i,w} S$. One might also consider conditions that connect authority relations for different propositions, e.g., $\left(S \preceq^P_{i,w} S' \text{ and } S \preceq^Q_{i,w} S'\right) \Rightarrow S \preceq^{P\cap Q}_{i,w} S'$ (cf. Liau 2003, p. 37). We leave such questions about the "logic of authority" aside in what follows.

Definition 4.4. The truth definition for the static part of DTL is:

- $\mathcal{M}, w \vDash U\varphi$ iff for all $v \in W$: $\mathcal{M}, v \vDash \varphi$
- $\mathcal{M}, w \vDash \mathtt{rec}_i\varphi$ iff $[\![\varphi]\!] \in \mathsf{rec}_i(w)$
- $\mathcal{M}, w \vDash S \preceq^\varphi_i S'$ iff $S \preceq^{[\![\varphi]\!]}_{i,w} S'$

4.3 Defining Trust with Record and Authority

In DTL it is trust that determines whether one agent accepts the testimony of another. In this section we briefly discuss how trust may be defined in terms of the testimonial record and authority relations. Since the aim of this paper is to provide a general framework for modeling the role of trust in testimony, we will not fix "the" definition of trust in DTL. Rather, we will indicate one possibility among many. The question of how to best define trust in terms of the record and authority deserves a full treatment on its own (cf. Cantwell 1998).

Suppose that we have defined what it is for an agent i to have testified *for* a proposition, $\mathsf{for}_i\varphi$, and *against* a proposition, $\mathsf{ags}_i\varphi$. Whether an agent is for or against a proposition will depend on which propositions she has testified to, according to the testimonial record, but let us leave the definitions open for a moment. Where $\boldsymbol{A}$ is the set of sequences of sets $\langle X_1, X_2, X_3\rangle$ such that $\{X_n \mid X_n \neq \emptyset\}$ is a *partition* of Agt, we might define a *trust formula* $T_{ji}\varphi$, read "j trusts the testimony of i on φ," as follows:

$$\alpha \wedge \bigvee_{\langle X,Y,Z\rangle\in \boldsymbol{A}} \left(\bigwedge_{x\in X} K_j\mathtt{for}_x\varphi \wedge \bigwedge_{y\in Y} K_j\mathtt{ags}_y\varphi \wedge \bigwedge_{z\in Z} \neg\left(K_j\mathtt{for}_z\varphi \vee K_j\mathtt{ags}_z\varphi\right) \wedge \beta\right)$$

where α and β are parameters. The parameter α determines the extent to which j must judge i doxastically reliable in order for j to trust i's testimony. The discussion in Section 1.2 suggest that only a minimal assumption of i's doxastic reliability is necessary, such as $\alpha := \hat{K}_j B_i\varphi \rightarrow \hat{K}_j(B_i\varphi \wedge \varphi)$. The parameter β determines the kind of trust defined. For example, for *weak trust* set $\beta := Y \setminus \{i\} \prec^\varphi_j X \cup \{i\}$. An agent j *weakly trusts* i's testimony on φ just in case if i were to join the group of agents whom j knows are *for* φ (and leave the group of agents whom j knows are *against*), then j would consider the group of agents who are for φ to be more authoritative on φ than the group of agents who are against. Stronger definitions of trust, which require additionally that j take i to be individually authoritative on φ, are also possible.

In addition to $T_{ji}\varphi$, we can define distrust $D_{ji}\varphi$, read"*j distrusts* the testimony of i on φ," and $A_{ji}\varphi$, read "j is *ambivalent* about i's testimony on φ," by changing α and β in the trust formula. For distrust set $\alpha := \hat{K}_j B_i \varphi \rightarrow \hat{K}_j (B_i \varphi \wedge \neg\varphi)$ and $\beta := X \cup \{i\} \prec_j^{\varphi} Y \setminus \{i\}$. For ambivalence set $\alpha := \hat{K}_j B_i \varphi \rightarrow \left(\hat{K}_j (B_i \varphi \wedge \varphi) \wedge \hat{K}_j (B_i \varphi \wedge \neg\varphi)\right)$ and $\beta := Y \setminus \{i\} \approx_j^{\varphi} X \cup \{i\} \wedge \neg \left(X \cup \{i\} \approx_j^{\varphi} \emptyset\right)$. The intuitions behind these definitions are easily grasped by analogy with the definition of trust. Note that the sense of "distrust" here is distrust in the judgment of i, which does not imply that j doubts the sincerity of i in testifying on φ.

Turning to $\mathsf{for}_i\varphi$ and $\mathsf{ags}_i\varphi$, the simplest option is to define $\mathsf{for}_i\varphi := \mathsf{rec}_i\varphi \wedge \neg\mathsf{rec}_i\neg\varphi$ and $\mathsf{ags}_i\varphi := \mathsf{rec}_i\neg\varphi \wedge \neg\mathsf{rec}_i\varphi$. Defining trust in this way, which we might call *narrow trust*, is sufficient for modeling "single-issue" testimonial sequences in which agents either testify for a single proposition or for its complement (or pass). For modeling sequences in which agents testify on multiple, related propositions, we may wish to consider a *wide trust* that depends on the authority not only of those who have testified that φ and those who have testified that $\neg\varphi$, but also of those who have testified that $\psi \wedge \varphi$, or ψ and $\psi \rightarrow \neg\varphi$, etc. In the interest of space, we leave aside a discussion of such wide trust here.

In addition to changing the definition of $\mathsf{for}_i\varphi$ and $\mathsf{ags}_i\varphi$, we might change the structure of the trust formula $T_{ji}\varphi$ itself in various ways. First, the operator B_j might be used instead of K_j in $T_{ji\varphi}$, so that j considers the authority not only of those whom she knows to be for or against φ, but also of those whom she believes to be for or against φ. However, for public testimony this makes little difference, since (as we will see in the next section) whenever an agent testifies that φ, all other agents come to *know* that the testifier is on the record for φ. Second, as we have defined $T_{ji}\varphi$, whether j trust i on φ depends only on the authority of i and the authority of those *on the record* for various propositions. Other policies are possible. For example, j may choose not to consider k's authority in favor of φ if although k testified that φ, j believes that k no longer believes φ.[8] Third and finally, as we have defined $T_{ji}\varphi$, j does not count his

[8]By counting the authority of those agents who have testified that φ but who (j believes) no longer believe φ, one assumes that j judges the testimonial reliability of k in terms of how reliable k's sincere testimony has been in the past, even in cases where (j believes) k later gave up the belief that the testimony expressed. We could distinguish such pure testimonial reliability from a hybrid

own authority in favor of or against φ. Yet we may wish to define $T_{ji}\varphi$ in such a way that whether j trusts i on φ depends on, among other things, j's current belief concerning φ and j's view of his own authority relative to others.

In what follows, we will assume that a definition of trust is given in terms of the testimonial record and authority relations. Our proposal for the dynamics of testimony does not depend on the details of the trust definition.

4.4 The Dynamics of Testimony

In this section, we define model transformations induced by the action of public testimony. The motivating idea is that if an agent j trusts another agent i on φ, then after i testifies that φ, j should come to believe φ. Moreover, if there is a *presumption of sincerity* in testimony, j should also come to believe that i believes φ. We begin by showing how to model the first part, j's belief revision concerning φ, by itself. We then show how to add the presumption of sincerity.

Throughout we let $\uparrow_i$ and $\downarrow_i$ stand for arbitrary belief upgrade and suspension operators respectively. Hence when we define model transformations below, we will actually be defining classes of model transformations, members of which differ with respect to the particular belief change operations used.

4.4.1 From Global to Local Belief Upgrades

In the simplest model of testimony, after an agent i testifies that φ, each agent j who trusts i's testimony on φ performs the belief upgrade $\uparrow_j \varphi$. Each agent who distrust i's testimony performs the upgrade $\uparrow_j \neg\varphi$, and each agent who is ambivalent about i's testimony performs the suspension $\downarrow_j \varphi$. The problem with this proposal is that $T_{ji}\varphi$ may be true at some states in the model and false at others—similarly for $D_{ji}\varphi$ and $A_{ji}\varphi$—reflecting other agents' uncertainty about j's attitude toward i. Suppose that $T_{ji}\varphi$ is true at the state at which we are evaluating formulas, so after i testifies that φ, j upgrades with $\uparrow_j \varphi$. Since belief upgrades work *globally* on the model, after the upgrade, $B_j\varphi$ may come to be true at a state in the model at which $T_{ji}\varphi$ is *false*. Hence agents who believe that j does *not* trust i on φ may nonetheless come to believe that j believes φ after i's testimony, a counterintuitive result.

The solution to this problem is to make global belief upgrades act locally. There is a technique for doing so, given by the following definition and lemma.

testimonial-doxastic reliability, judged by the reliability of an agent's testimony in just those cases in which the agent retained the belief that the earlier testimony expressed.

Definition 4.5. A formula φ is *introspectible* for an agent i in a model $\mathcal{M}$ iff for every information cell $C \subseteq W$ for i, $[\![\varphi]\!] \cap C \neq \emptyset \Rightarrow C \subseteq [\![\varphi]\!]$.

The introspectible formulas for i in $\mathcal{M}$ are those φ such that at any state in $\mathcal{M}$, if i considers φ possible, then i knows φ. Examples (for any model) include belief and knowledge formulas $B_i\psi$ and $K_i\psi$ and trust formulas $T_{ik}\psi$.

Lemma 1 (Localization)**.** Let $\mathcal{M}$ be a multi-agent plausibility model, $\psi_1, .., \psi_n$ a sequence of formulas, and $\varphi_1, ..., \varphi_n$ a sequence of introspectible formulas for i in $\mathcal{M}$ such that $[\![\varphi_j]\!] \cap [\![\varphi_k]\!] = \emptyset$ for $j \neq k$. Then there is a formula χ such that for every information cell $C \subseteq W$ for i with $[\![\varphi_k]\!] \cap C \neq \emptyset$, it holds that $\leq_i^{\mathcal{M}\Uparrow_i\chi} \upharpoonright C = \leq_i^{\mathcal{M}\Uparrow_i\psi_k} \upharpoonright C$.

Hence the effect of a single upgrade with the formula χ is that each comparability class containing a point that satisfies one of the φ_k is reordered locally just as it would be by a (global) upgrade by ψ_k. Intuitively, χ "localizes" revision by ψ_k to those parts of the model "targeted" by φ_k.

Proof. Take $\chi := \bigvee_{1 \leq k \leq n} (\varphi_k \wedge \psi_k)$. Suppose C is an information cell for i with $[\![\varphi_k]\!] \cap C \neq \emptyset$. Then since φ_k is introspectible, $C \subseteq [\![\varphi_k]\!]$. It follows by the assumption of the lemma that $[\![\varphi_j]\!] \cap C = \emptyset$ for $j \neq k$. Hence $[\![\chi]\!] \cap C = [\![\varphi_k \wedge \psi_k]\!] \cap C = [\![\varphi_k]\!] \cap [\![\psi_k]\!] \cap C$. Given $C \subseteq [\![\varphi_k]\!]$ we also have $[\![\varphi_k]\!] \cap [\![\psi_k]\!] \cap C = [\![\psi_k]\!] \cap C$. Therefore $[\![\chi]\!] \cap C = [\![\psi_k]\!] \cap C$, which gives $\leq_i^{\mathcal{M}\Uparrow_i\chi} \upharpoonright C = \leq_i^{\mathcal{M}\Uparrow_i\psi_k} \upharpoonright C$. □

The following application of the Localization Lemma shows the utility of making a number of different local changes to a model with one global upgrade.

Definition 4.6 (Testimonial Upgrade)**.** Let $\Updownarrow_j^i \varphi$ be the operation that performs the following sequence of relation changes:

$$\Uparrow_j \left(T_{ji}\varphi \wedge \varphi\right) \vee \left(D_{ji}\varphi \wedge \neg\varphi\right), \downarrow_j A_{ji}\varphi \wedge \varphi$$

We propose that after i testifies that φ, j's plausibility ordering should change according to $\Updownarrow_j^i \varphi$.[9] To understand the effect of the operation $\Updownarrow_j^i \varphi$, note that each state in a testimonial model satisfies at most one of $T_{ji}\varphi$, $D_{ji}\varphi$, and $A_{ji}\varphi$,

[9]Note that the definition of $\Updownarrow_j^i \varphi$ is not supposed to reflect j's *mental representation* of the information given by i's testimony; j need not think of the information received from i's testimony as "either I trust j on φ and φ, or...." The point is rather to identify a change to j's plausibility ordering that produces an intuitively correct effect on the testimonial model, for it is the testimonial model that reflects the epistemic situation of j after i's testimony.

as these are mutually exclusive. Moreover, in a *legal* model, the states in any given information cell for j must all agree on which one of these formulas they satisfy, by the fourth condition of legality. Hence in any given information cell for j in which $T_{ji}\varphi$ is true, the above upgrade sequence will have the same effect *as if* j *upgraded with* $\uparrow_j \varphi$ *alone*, since all the states in the information cell satisfy $T_{ji}\varphi$ and none satisfy $D_{ji}\varphi$ or $A_{ji}\varphi$. Similarly, in any given information cell for j in which $D_{ji}\varphi$ is true, the upgrades will have the same effect as if j upgraded with $\uparrow_j \neg\varphi$ alone, and so on. The localization technique allows us in effect to do *different belief revision* for j in *different parts of the model*, depending on j's attitude toward i (trust, distrust, etc.) in different parts of the model. Hence if another agent k is uncertain about whether or not j trusts i, then after i's testimony, k will be uncertain about which belief revision j performed. We illustrate this phenomenon in the course of modeling the bandwagon effect in Section 5.

We are now ready to define the model transformation for public testimony with *no presumption of sincerity* (NPS).

Definition 4.7 (Public Testimony with NPS)**.** Given a testimonial model $\mathcal{M} = \langle W, \leq, V, \mathsf{rec}, \preceq \rangle$, the model $\mathcal{M}!_i\varphi = \left\langle W, \leq^{!_i\varphi}, V, \mathsf{rec}^{!_i\varphi}, \preceq \right\rangle$ is defined as follows: $\mathsf{rec}_j^{!_i\varphi}(w) = \mathsf{rec}_j(w)$ for $j \neq i$, $\mathsf{rec}_i^{!_i\varphi}(w) = (\mathsf{rec}_i(w) \setminus \{W \setminus [\![\varphi]\!]\}) \cup \{[\![\varphi]\!]\}$, and $\leq^{!_i\varphi}$ is obtained from $\leq$ by the sequence of operations $\updownarrow^i_{j \in \mathsf{Agt} \setminus \{i\}} \varphi$.

The definition states that when i testifies that φ, we put i on the record *at each state* w for $[\![\varphi]\!]$, reflecting the public nature of the testimony. Moreover, we take i off the record for $W \setminus [\![\varphi]\!]$, reflecting the assumption that by testifying that φ, i is implicitly retracting past testimony against φ. We could also define a general *retraction* operation allowing agents to retract past testimony for some ψ without having to go on the record for $W \setminus [\![\psi]\!]$, but we will not do so here.

The notation $\updownarrow^i_{j \in \mathsf{Agt} \setminus \{i\}} \varphi$ indicates that every agent j other than i performs the individual testimonial upgrade $\updownarrow^i_j \varphi$. (Since the agents are not upgrading by doxastic formulas, the order in which the upgrades occur does not matter.) Note that if $\mathcal{M}$ is legal, then the updated model $\mathcal{M}!_i\varphi$ is also legal, since belief upgrade and suspension do not change the relations $\preceq_j$ or $\preceq^P_{j,w}$ and since the operation $!_i$ makes the same changes to $\mathsf{rec}_j(w)$ and $\mathsf{rec}_j(v)$ for $w \sim_j v$.[10]

[10] Since it is permitted in a testimonial model for $\mathsf{rec}_j(w)$ and $\mathsf{rec}_j(v)$ to differ for $w \nsim_j v$, we might also consider testimony that is less than fully public, which leaves some agents uncertain about the content of the testimony. This would require more complex model transformations, and private testimony "behind closed doors" may require changes to the definition of a testimonial model (cf. van Ditmarsch et al. 2008, Sec. 6.9).

Definition 4.8. The truth definition for the testimony operator is:

- $\mathcal{M}, w \models [!_i\varphi]\,\psi$ iff $\mathcal{M}, w \models \texttt{pre}$ implies $\mathcal{M}!_i\varphi, w \models \psi$

where the precondition $\texttt{pre} := B_i\varphi$ with the *assumption of sincerity* (AS) and $\texttt{pre} := \top$ without. In addition to the box modality $[!_i\varphi]$, we can define a dual diamond modality $\langle !_i\varphi \rangle$ by replacing 'implies' by 'and' in the definition above.

4.4.2 Modeling the Presumption of Sincerity

Given a *presumption of sincerity* (PS) in testimony, if j trusts i on φ and i testifies that φ, j should come to believe not only that φ but also that i believes φ. Our first question is whether we should model j's belief revision in a single step, in which j performs the upgrade $\uparrow_j B_i\varphi \wedge \varphi$, or in two steps, in which j first performs $\uparrow_j B_i\varphi$ and then $\uparrow_j \varphi$ or *vice versa*. Technically, we could model the revision either way. Conceptually, it seems preferable to model it in two steps.

Suppose that agent j receives testimony from i that φ, followed by testimony from k that $\neg\varphi$. While j considers i authoritative on φ, j considers k still more authoritative. However, i does not consider k authoritative at all, and j knows this. Intuitively, after the testimonies j should believe that i believes φ, that k believes $\neg\varphi$, and that $\neg\varphi$. However, if we model j's first belief revision with the single upgrade $\uparrow_j B_i\varphi \wedge \varphi$, we may not obtain the desired result. Instead, after k's testimony j may lose his belief that $B_i\varphi$, formed given his presumption of i's sincerity, even though j *knows* that i does not trust k on $\neg\varphi$. Figure 1[11] illustrates how this counterintuitive result may occur. To simplify, we do not draw arrows for i or represent k in the model at all. However, a more complex model with k represented would give the same result. Note that the loss of j's belief in $B_i\varphi$ after the second upgrade is independent of the kind of belief upgrades used.

If we wish to model j's belief revision after i's testimony in two steps, the obvious candidates are the sequences of upgrades $\uparrow_j B_i\varphi, \uparrow_j \varphi$ and $\uparrow_j \varphi, \uparrow_j B_i\varphi$. Both sequences avoid the problem of j too easily losing his belief that $B_i\varphi$.[12] However, both sequences also violate an intuitive constraint on j's

[11] In the following Figures, circles represent states and lines represent plausibility orderings, labelled for each agent, with arrows pointing toward more plausible states and arrowless lines indicating equi-plausibility. Every state is equi-plausible with itself for each agent, but we omit the reflexive loops, as well as arrows implied by transitivity. Atomic sentences true at a state are indicated inside the circle representing the state; all other atomic sentences are false at the state.

[12] Unless $\uparrow_j B_i\varphi$ is a conservative upgrade, in which case j may fail to believe $B_i\varphi$ after $\uparrow_j B_i\varphi, \uparrow_j \varphi$ (even if j considers it possible that $B_i\varphi \wedge \varphi$) or one upgrade after $\uparrow_j \varphi, \uparrow_j B_i\varphi$, if as before k testifies that

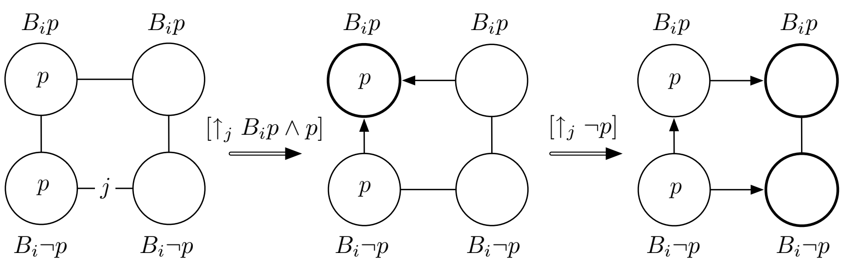

Figure 1: A failure to preserve belief in the testifier's sincerity

belief revision in response to i's testimony, namely that the revision should not promote any states satisfying $\varphi \wedge \neg B_i\varphi$ to become more plausible. The only reason j is now upgrading with φ is that j took i to have sincerely testified for φ, so it does not make sense for j to promote a φ-state at which i does not believe φ. Yet the sequences $\uparrow_j B_i\varphi, \uparrow_j \varphi$ and $\uparrow_j \varphi, \uparrow_j B_i\varphi$ can clearly have this effect.

We can respect the constraint on state promotion as follows. If j trusts i's testimony on φ, then we use the sequence $\uparrow_j B_i\varphi, \uparrow_j B_i\varphi \wedge \varphi$ to model j's belief revision after i's testimony that φ. If j distrusts i's testimony, we use $\uparrow_j B_i\varphi, \uparrow_j B_i\varphi \wedge \neg\varphi$, and if j is ambivalent about i's testimony, we use $\uparrow_j B_i\varphi, \downarrow_j \{B_i\varphi \wedge \varphi, B_i\varphi \wedge \neg\varphi\}$. (Note here the essential use of multi-upgrade.) Otherwise j performs only the upgrade $\uparrow B_i\varphi$, in which case whether j comes to believe φ will depend on whatever beliefs j has about i's doxastic reliability.

Definition 4.9 (Testimonial Upgrade with PS)**.** Let $\gamma := B_i\varphi \wedge \varphi$, $\gamma' := B_i\varphi \wedge \neg\varphi$. The operation $\updownarrow\!\updownarrow^i_j \varphi$ performs the following sequence of relation changes:

$$\uparrow_j B_i\varphi, \uparrow_j \left(T_{ji}\varphi \wedge \gamma\right) \vee \left(D_{ji}\varphi \wedge \gamma'\right), \downarrow_j \left\{A_{ji}\varphi \wedge \gamma, A_{ji}\varphi \wedge \gamma'\right\}$$

Definition 4.10 (Public Testimony with PS)**.** Given a testimonial model $\mathcal{M} = \langle W, \leq, V, \mathsf{rec}, \preceq \rangle$, the model $\mathcal{M}!_i\varphi = \left\langle W, \leq^{!_i\varphi}, V, \mathsf{rec}^{!_i\varphi}, \preceq \right\rangle$ is defined as follows: $\mathsf{rec}_j^{!_i\varphi}(w)$ is determined as in Definition 4.7, and $\leq^{!_i\varphi}$ is obtained from $\leq$ by the sequence of operations $\updownarrow\!\updownarrow^i_{j \in \mathsf{Agt}\setminus\{i\}} \varphi$.

The truth definition for testimony with PS is the same as with NPS.

$\neg\varphi$ and j performs the upgrade $\uparrow_j \neg\varphi$ (even if j considers it possible that $B_i\varphi \wedge \neg\varphi$). Conservative upgrade has been criticized for similar failures to preserve belief revisions (Booth and Meyer 2006).

4.5 Basic DTL Validities

Given the semantics developed in the previous section, a number of interesting validities holds in DTL for the class of legal models. The validities below do not depend on the definitions of $T_{ji}\varphi$, $D_{ji}\varphi$, and $A_{ji}\varphi$, except that the definitions guarantee the validity of $T_{ji}\varphi \to \alpha$ for the parameter α from Section 4.3, that is,

(0) $\vDash T_{ji}\varphi \to \left(\hat{K}_j B_i\varphi \to \hat{K}_j (B_i\varphi \land \varphi)\right)$

and likewise for $D_{ji}\varphi$ and $A_{ji}\varphi$ with their respective values of α.

We have organized selected validities according to the questions about testimony with which we began in Section 1.1. The semantic conditions for which each formula is valid are indicated in parentheses to the right of the formula.

What do agents in i's audience come to believe about φ?

(1.1) $\vDash T_{ji}\varphi \to [!_i\varphi]\, B_j\varphi$ (NPS or AS)
(1.2) $\vDash D_{ji}\varphi \to [!_i\varphi]\, B_j\neg\varphi$ (NPS or AS)
(1.3) $\vDash A_{ji}\varphi \to [!_i\varphi]\, (\neg B_j\varphi \land \neg B_j\neg\varphi)$ (NPS or AS)

The semantic requirement of NPS or AS indicates that if we have the presumption of sincerity, then we must also have the assumption of sincerity. The reason (1.1) holds with NPS is that if j trusts i on φ, then by the third condition on legal models there is a state accessible for j from the current state that satisfies φ. Therefore, when i testifies that φ and j upgrades with φ, j's upgrade will be successful. The reason (1.1) holds for AS is that if i believes φ, then j considers it possible that $B_i\varphi$ (given reflexivity for $\leq_j$), and if moreover j trusts i on φ, then j also considers it possible that $B_i\varphi \land \varphi$ by (0) above. Hence when i testifies that φ and j upgrades with $B_i\varphi \land \varphi$ for PS or with φ for NPS, j's upgrade will be successful. The explanations for (1.2) and (1.3) are similar. Note that if instead of (0) we had (0′) $\vDash T_{ji}\varphi \to \hat{K}_j (B_i\varphi \land \varphi)$, then (1.1) would always hold (and similarly for (1.2)-(1.3) with the analogous changes for $D_{ji}\varphi$ and $A_{ji}\varphi$).

What do agents in i's audience come to believe about i's beliefs?

(2.1) $\vDash [!_i\varphi]\, B_j B_i\varphi$ (AS, PS)
(2.2) $\vDash \hat{K}_j B_i\varphi \to \langle !_i\varphi\rangle\, B_j B_i\varphi$ (NAS, PS)

and similarly for $D_{ji}\varphi$ and $A_{ji}\varphi$.

The reason (2.1) holds for AS and PS is that given PS, after i testifies that φ, j will upgrade with $B_i\varphi$; but given AS, i believes φ, so j considers $B_i\varphi$ possible (by reflexivity for $\leq_j$ again), which guarantees the success of j's upgrade. Without AS, it must be built into the antecedent of (2.2) that j considers $B_i\varphi$ possible.

What do agents come to believe about the beliefs of other agents in i's audience?

(3.1) $\vDash K_k T_{ji}\varphi \rightarrow [!_i\varphi]\, K_k B_j\varphi$ (NPS or AS)
(3.2) $\vDash [!_i\varphi]\, B_k T_{ji}\varphi \rightarrow [!_i\varphi]\, B_k B_j\varphi$ (NPS or AS)

and similarly for $D_{ji}\varphi$ and $A_{ji}\varphi$.

The reason (3.1) holds is that if k knows that j trusts i on φ, then no matter how k revises her beliefs in response to i's testimony, k will still know that j trusts i on φ after i testifies that φ, i.e., $\vDash K_k T_{ji}\varphi \rightarrow [!_i\varphi]\, K_k T_{ji}\varphi$. By (1.1) every state satisfying $T_{ji}\varphi$ will also satisfy $B_j\varphi$ after i testifies that φ, so k will know $B_j\varphi$ after i testifies that φ. The explanation of (3.2) is similar. The reason that (3.1) does not hold with B_k in place of K_k is that while k may believe that j trusts i on φ before i testifies that φ, after i testifies that φ and k revises her beliefs accordingly, k may no longer believe that j trusts i on φ, i.e., $\nvDash B_k T_{ji}\varphi \rightarrow [!_i\varphi]\, B_k T_{ji}\varphi$.[13]

5 Application: Information Cascades

In this section we give an application of DTL in modeling the example of an information cascade introduced in Section 1, Sorensen's (1984) epistemic bandwagon effect. Given the intended readings of the DTL formulas, let the initial premises about the three experts in Sorensen's scenario be:

$$\bigwedge_{i \in \mathsf{Agt}} \left(\neg \mathsf{rec}_i p \wedge \neg \mathsf{rec}_i \neg p\right)$$

1. $B_1 p$ $\quad \emptyset \approx_1^p \{2\} \approx_1^p \{3\} <_1^p \{2,3\}$
2. $\neg B_2 p \wedge \neg B_2 \neg p$ $\quad \emptyset <_2^p \{1\} \approx_2^p \{3\} <_2^p \{1,3\}$
3. $B_3 \neg p$ $\quad \emptyset \approx_3^p \{1\} \approx_3^p \{2\} <_3^p \{1,2\}$

[13] Could such a situation plausibly come about? We will suggest the structure that such a situation may have, leaving it to the reader to provide a concrete example. Suppose that k believes that j trusts i on φ, but k also believes that i does not believe φ. Further suppose that k has a conditional belief to the effect that if i *does* believe φ, then another agent l believes $\neg\varphi$. Finally, suppose that k has a conditional belief to the effect that if l believes $\neg\varphi$, then j believes that l believes $\neg\varphi$ and j does not trust anyone on φ. Given these assumptions, when i testifies that φ, k will come to believe that l believes $\neg\varphi$ and hence that j does not trust i on φ.

Let us assume that $[!_i\varphi]$ is the operator for testimony with the presumption of sincerity, defined as in Definition 4.10, where in Definition 4.9 $T_{ji}\varphi$ is the narrow, weak trust of Section 4.3 and $\Uparrow_i \varphi$ is any of the upgrade operations given in Section 3. Under these assumption we define a new abbreviation $\langle ?_i\varphi\rangle\, \psi$, read "after i testifies with her opinion on φ, ψ is the case" as follows:

$$(\neg B_i\varphi \wedge \neg B_i\neg\varphi \wedge \langle !_i\top\rangle\, \psi) \vee \langle !_i\varphi\rangle\, \psi \vee \langle !_i\neg\varphi\rangle\, \psi$$

With this definition, the following formulas represent the outcomes of Sorensen's two testimonial sequences:

$$\langle ?_1p\rangle \langle ?_2p\rangle \langle ?_3p\rangle\, B_1p \wedge B_2p \wedge B_3p$$

$$\langle ?_3p\rangle \langle ?_2p\rangle \langle ?_1p\rangle\, B_1\neg p \wedge B_2\neg p \wedge B_3\neg p$$

These formulas are derivable in DTL from the premises above under the stated assumptions, but we will not give the derivation here. Instead, we will show how to model the bandwagon effect semantically.

Model $\mathcal{M}_0$ in Figure 2 represents some of the relevant first and second-order beliefs of the three experts in the initial situation, focusing on the perspective of expert 3. We assume that the authority relations given above for the three agents hold everywhere in $\mathcal{M}_0$, except that in all of the states on the right side of the model, 2 does not consider 1 authoritative on p. The shaded circle on the left represents the "actual state." At the actual state, expert 1 believes p, expert 3 believes $\neg p$, and expert 2 is undecided. There is also information about each expert's beliefs about the beliefs and authority relations of the others, which we have added to Sorensen's original scenario. For example, 3 knows that 1 believes p, while 2 is uncertain about what 1 believes, and 3 knows this about 2. Finally, 3 is uncertain about whether or not 2 considers 1 authoritative on p.

When expert 1 testifies that p, two things happen. First, experts 2 and 3 both upgrade with B_1p, given the presumption of sincerity. Since 3 already knows that 1 believes p, nothing changes for 3. But 2's plausibility ordering does change, reflected in the transition from $\mathcal{M}_0$ to $\mathcal{M}_1$ in Figure 3. Second, since on the left side of the model 2 considers 1 authoritative on p, and since no one else has testified to the contrary, on the left side of the model 2 *trusts* 1 on p. Hence when 2 performs the upgrade $\Uparrow_2 (T_{21}p \wedge B_1p \wedge p) \vee (D_{21}\varphi \wedge B_1p \wedge \neg p)$, this upgrade changes the model as if the upgrade $B_1p \wedge p$ occurred on the left side of the model only, as reflected in the transition from $\mathcal{M}_1$ to $\mathcal{M}_2$. (The testimony operation $!_1p$ transforms $\mathcal{M}_0$ to $\mathcal{M}_2$ in one step, but we have broken up the transformation for the purpose of explanation.) Note that 3's initial

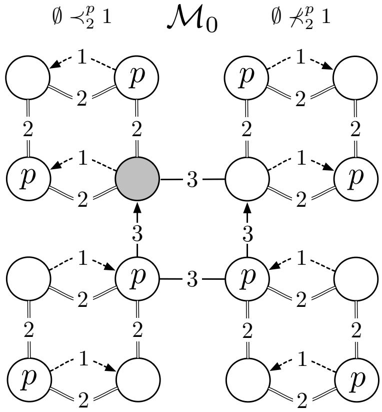

Figure 2: The initial epistemic situation of the three experts

uncertainty in $\mathcal{M}_0$ about whether 2 *trusts* 1 on p leads to uncertainty for 3 in $\mathcal{M}_2$ about whether 2 *believes* p after 1's testimony. By assumption, 3 does not trust 1 on p, so there are no further changes to 3's ordering.

Having been convinced of p by the testimony of 1, 2 now testifies that p. Once again, two things happen. First, 1 and 3 upgrade with B_2p, given the presumption of sincerity. Nothing changes for 1, but 3's plausibility ordering does change, as reflected in the transition from $\mathcal{M}_2$ to $\mathcal{M}_3$. Not only does 3 come to believe that 2 believes p, but also 3 comes to believe *that* 2 *trusts* 1 *on* p, illustrating how an agent can learn from another agent's testimony whom the testifier trusts. Second, since 3 considers 1 and 2 together *jointly* authoritative on p, 3 now trusts 2 on p. Hence when 3 upgrades with $\Uparrow_3 (T_{32}p \wedge B_2p \wedge p) \vee (D_{32}\varphi \wedge B_2p \wedge \neg p)$, 3 comes to believe p, as reflected in the transition from $\mathcal{M}_3$ to $\mathcal{M}_4$. At this point, the epistemic bandwagon effect has occurred. All agents (falsely) believe p and believe that the others believe p.

Our model not only reflects how bandwagons start, but also suggest one way to stop them. Since expert 3 knew that expert 2 was initially undecided about p, and since the only informational event to occur before 2's testimony was 1's testimony, when 2 testified that p, 3 came to believe that 2 trusts 1 on p. If 3 were more sophisticated, 3 might have reasoned that since 1's testimony *influenced* 2 on p, the joint authority of 1 and 2 on p should not be greater than the authority of 1 alone, in which case 3 would not have come to believe p after the testimony of 2. To capture this reasoning formally, we would have to represent

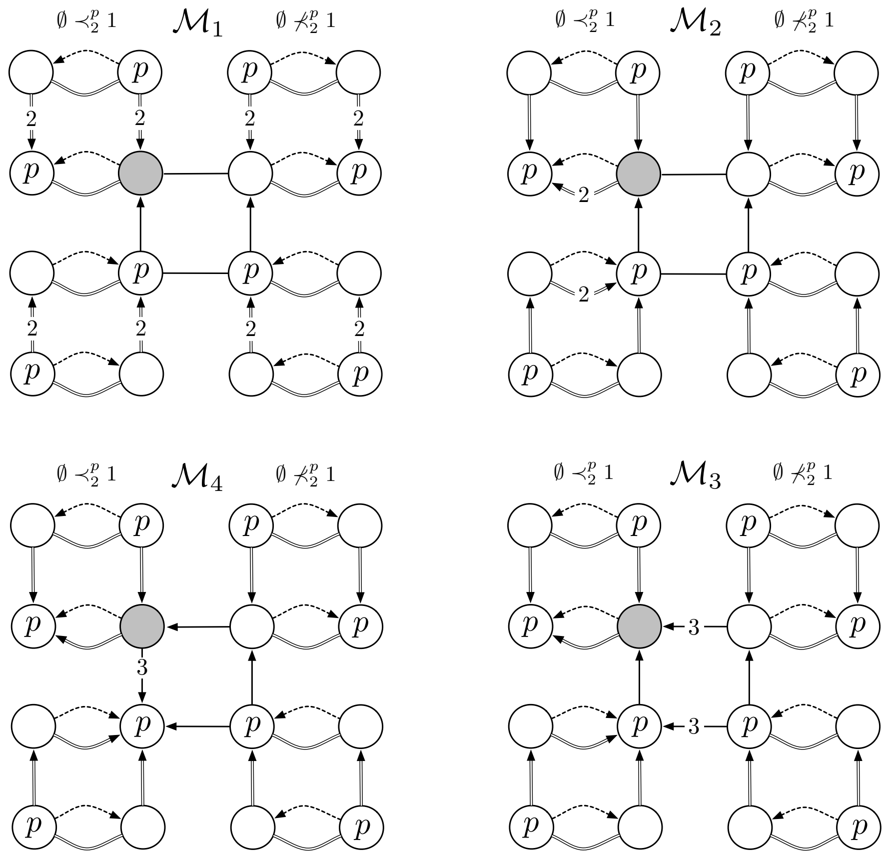

Figure 3: An epistemic bandwagon effect (clockwise from upper left)

influence explicitly in our model, and we would have to provide a mechanism by which agents can change their authority relations. These additions are beyond the scope of this paper, but they suggest that it may be possible to do more than modeling the bad news about bandwagons.

6 Conclusion

We have proposed a dynamic testimonial logic (DTL) to model belief change over sequences of testimony among agents with different dispositions to trust each other as information sources. DTL adds to standard DEL the semantic structures of *testimonial records* and *authority relations* and the dynamic action of *testimony*. In the framework of DTL, we have shown how to define epistemic trust in terms of the record and authority and how to use the technique of "localizing" belief upgrades to simultaneously do different belief revisions for an agent in different parts of a model, determined by the agent's attitude (trust, distrust, or ambivalence) in different parts of the model toward an information source. We have also shown how to capture the presumption of sincerity in testimony with the right choice of belief upgrades. Finally, our DTL model of the epistemic bandwagon showed, first, how an agent's uncertainty about whom another agent trusts can lead to uncertainty about what the other agent believes and, second, how an agent may learn from testimony about whom a testifier trusts. The Appendix contains a complete axiomatization for DTL.

For future work, there are a number of possible extensions to the framework of DTL. These include working out a more sophisticated definition of trust, adding trust and testimony on doxastic formulas, representing testimony that is not fully public, and modeling agents' reasoning about how testimony has influenced the beliefs of others. In an extended version of DTL, we might model not only how bandwagons start, but also how to stop them.

Acknowledgements I wish to thank Johan van Benthem, Tomohiro Hoshi, Thomas Icard, and Eric Pacuit for helpful feedback on this paper. I also benefited from discussion of this work at the 2009 Logical Dynamics Workshop at Stanford University and at the 2009 LORI-II Workshop at Southwest University in Chongqing, China. I thank the organizers and participants at those events.

A Appendix

Theorem 1. *Together with an axiomatization of CDL (plus the universal modality for propositional formulas) and reduction axioms for a pair of belief upgrade and suspension operators, the following axiom system for DTL is sound and complete for the class of legal testimonial models.* The static axioms are the following:

(R1) $U(\varphi \leftrightarrow \psi) \rightarrow (\mathtt{rec}_i\varphi \leftrightarrow \mathtt{rec}_i\psi)$
(R2) $\mathtt{rec}_i\varphi \rightarrow K_i\mathtt{rec}_i\varphi$
(A1) $U(\varphi \leftrightarrow \psi) \rightarrow \left(S \preceq_i^{\varphi} S' \leftrightarrow S \preceq_i^{\psi} S'\right)$
(A2) $S \preceq_i^{\varphi} S$
(A3) $\left(S \preceq_i^{\varphi} S' \wedge S' \preceq_i^{\varphi} S''\right) \rightarrow S \preceq_i^{\varphi} S''$
(A4) $S \preceq_i^{\varphi} S' \leftrightarrow S \preceq_i^{\neg\varphi} S'$
(A5) $S \preceq_i^{\varphi} S' \rightarrow \hat{K}_i\left(\varphi \wedge \bigwedge_{j\in S}\neg K_j\varphi\right)$
(A6) $S \preceq_i^{\varphi} S' \rightarrow K_i S \preceq_i^{\varphi} S'$

The additional reduction axioms for belief upgrade and suspension in DTL are:

(B1) $[\pi, \varphi]\,\alpha \leftrightarrow \alpha$ for $\pi := \uparrow_i, \downarrow_i$ and $\alpha := \mathtt{rec}_j\psi, S \preceq_j^{\psi} S', U\psi$

The reduction axioms for the testimony operator with AS and PS are:

(T1) $[!_i\varphi]\,\mathtt{rec}_i\psi \leftrightarrow \left(B_i\varphi \rightarrow ((\mathtt{rec}_i\psi \wedge \neg U(\psi \leftrightarrow \neg\varphi)) \vee U(\psi \leftrightarrow \varphi))\right)$
(T2) $[!_i\varphi]\,\mathtt{rec}_j\psi \leftrightarrow \left(B_i\varphi \rightarrow \mathtt{rec}_j\psi\right)$ for $j \neq i$
(T3) $[!_i\varphi]\,\alpha \leftrightarrow (B_i\varphi \rightarrow \alpha)$ for $\alpha := p, S \preceq_i^{\psi} S', U\psi$
(T4) $[!_i\varphi]\,\neg\psi \leftrightarrow \left(B_i\varphi \rightarrow \neg\,[!_i\varphi]\,\psi\right)$
(T5) $[!_i\varphi]\,(\psi \wedge \psi') \leftrightarrow \left([!_i\varphi]\,\psi \wedge [!_i\varphi]\,\psi'\right)$
(T6) For θ and ψ that do not contain testimony operators:
$[!_i\varphi]\,B_j^{\theta}\psi \leftrightarrow \left(B_i\varphi \rightarrow r\left(\left[\updownarrow\updownarrow_{k\in X}^{i}\,\varphi\right]B_j^{\theta^*}\psi^*\right)\right)$

where $X = \left\{k \in \mathsf{Agt} \mid k \neq i \text{ and } k \text{ occurs in } B_j^{\theta}\psi\right\}$, $\left[\updownarrow\updownarrow_{k\in X}^{i}\,\varphi\right]$ abbreviates a string of upgrade and suspension operators corresponding to the relation changes in Definition 4.9, and r is the *reduction function* which, given a formula in the language of upgrade and suspension, uses the appropriate reduction axioms (as in van Benthem 2007) to return an equivalent, reduced formula in the language of CDL plus rec and $\preceq$. The formulas θ^* and ψ^* are obtained from θ and ψ respectively by replacing, for all α, each occurrence of $\mathtt{rec}_i\alpha$ in θ and ψ by an occurrence of $(\mathtt{rec}_i\alpha \wedge \neg U(\alpha \leftrightarrow \neg\varphi)) \vee U(\alpha \leftrightarrow \varphi)$.

We have given the reduction axioms for testimony with the *assumption of sincerity* and *presumption of sincerity*. For no assumption of sincerity, simply drop the precondition of $B_i\varphi$ in the conditionals of (T1) – (T4) and (T6). For no presumption of sincerity, replace $\left[\updownarrow\!\updownarrow^i_{k\in X}\ \varphi\right]$ in (T6) by the abbreviation $\left[\updownarrow^i_{k\in X}\ \varphi\right]$.

For CDL we assume the axioms of Baltag and Smets (2008, p. 37) or the equivalent system BRSIC of Board (2004, Sec. 3.3). For $U\varphi$ with φ propositional we take the S5 axioms plus $U\varphi \to K_i\varphi$ (cf. Blackburn et al. 2001, p. 415ff.).

Soundness. (R1) and (A1) hold in virtue of the truth definition for record and authority formulas. (R2) and (A6) hold in virtue of the fourth condition on legal models, (A2) and (A3) in virtue of the first, and (A4) and (A5) in virtue of the second and third respectively. (B1) holds because upgrades and suspensions do not change the record, authority relations, or universal propositional facts.

(T1) – (T6) hold by definition of the testimony operation on models. (T1) says that after i testifies that φ, i is on the record for ψ iff i was already on the record for ψ and ψ is not equivalent to $\neg\varphi$ in the model (for if they are equivalent, then i would have been taken off the record for ψ when she testified that φ) or ψ is equivalent to φ (in which case i was added to the record for ψ when she testified that φ). (T2) says that for agents other than i, the record does not change after i's testimony. (T3) reflects the fact that testimony does not change atomic facts, authority relations, or universal propositional facts. (T4) – (T5) give standard properties of dynamic operators.

(T6) captures the effect of i's testimony that φ on agents' beliefs, which by definition is determined by the sequence of testimonial upgrades $\updownarrow\!\updownarrow^i_{k\in X}\ \varphi$. Note that we only consider what happens to the beliefs of those agents whose symbols appear in $B^\theta_j\psi$, since the beliefs of others do not matter for evaluating the formula. Following the definition of the testimony operation, we do not change the testifier i's beliefs.

The reason for the change from θ and ψ to θ^* and ψ^* is that θ and ψ may contain a formula $\mathtt{rec}_i\alpha$, the truth value of which may change after i's testimony. Hence we use the same idea as in (T1) and express what must be true in the original model in order for $\mathtt{rec}_i\alpha$ to be true in the model updated by $!_i\varphi$. However, if θ or ψ contains testimony operators, the replacement of θ and ψ by θ^* and ψ^* may not achieve the correct result. For example, if ψ is $[!_i\neg\varphi]\,\mathtt{rec}_i\neg\varphi$, then ψ^* is $[!_i\neg\varphi]\,(\mathtt{rec}_i\neg\varphi \wedge \neg U(\neg\varphi \leftrightarrow \neg\varphi)) \vee U(\neg\varphi \leftrightarrow \varphi)$, and while ψ is valid, ψ^* is unsatisfiable. We avoid this problem by the restriction in (T6) that θ and ψ do not contain testimony operators.

Since we can reduce DTL formulas by applying the reduction axioms from the "inside out," eliminating testimony operators from subformulas first, the restriction on θ and ψ does not prevent us from reducing any DTL formula. Because θ and ψ are subformulas of $B^\theta_j\psi$, any testimony operators will be eliminated from them by the time we get to $[!_i\varphi]\, B_j^{r(\theta)} r(\psi)$.

Completeness. Using the reduction axioms (R1) – (T6), every formula of DTL with dynamic operators is reducible to an equivalent formula in the static part of the language. It therefore suffices to show completeness for the static part of DTL, which we will now sketch. Following the standard strategy, we show that if φ is not refutable by the axioms of DTL given above, then φ is satisfiable in a legal testimonial model. To produce the satisfying model we use the canonical model construction for CDL (Board 2004, Proof of Theorem 2, p. 77), but with two differences. The first difference is that although we construct maximally consistent sets (MCSs) from the subformulas of φ (which may now include formulas of the form $U\psi$, $\mathtt{rec}_i\psi$, and $S \preceq_i^{\psi} S'$) in the same way as for CDL, we do not take the domain of the canonical model to contain *all* such subformula-generated MCSs. Instead, where Γ_φ is a MCS containing φ and R_U is a relation on MCSs such that $\Sigma R_U \Delta$ iff $\{\psi \mid U\psi \in \Sigma\} = \{\psi \mid U\psi \in \Delta\}$, we take for the domain of the canonical model the set of subformula-generated MCSs Γ such that $\Gamma_\varphi R_U \Gamma$. The second difference is that we must also construct a testimonial record and authority relations in the canonical model for DTL.

Definition A.1. The canonical testimonial model based on φ is the model $\mathcal{M}_\varphi = (W, \leq, V, \mathsf{rec}, \preceq)$ with W defined as above, $\leq_i$ defined as for CDL by Board (2004), V defined as usual by $V(p) = \{\Gamma \in W \mid p \in \Gamma\}$, and $\mathsf{rec}_i(\Gamma)$ and $\preceq_{i,w}^{P}$ defined by:

- $\mathsf{rec}_i(\Gamma) = \{P \subseteq W \mid \text{there is an } \alpha \text{ with } \mathtt{rec}_i\alpha \in \Gamma \text{ and } [\![\alpha]\!] = P\}$
- $S \preceq_{i,\Gamma}^{P} S'$ iff there is an α with $S \preceq_i^{\alpha} S' \in \Gamma$ and $[\![\alpha]\!] = P$

Lemma 2 (Truth). $\mathcal{M}_\varphi, \Gamma \vDash \psi \Leftrightarrow \psi \in \Gamma$

Proof. By induction on ψ. We mention only the cases for $\mathtt{rec}_i\chi$ and $S \preceq_i^{\chi} S'$, leaving it as an exercise to the reader to check that Board's (2004) proof for the case of $B_i^{\psi}\chi$ works with our modified canonical model, given the axioms for U.

If $\mathcal{M}_\varphi, \Gamma \vDash \mathtt{rec}_i\chi$, then $[\![\chi]\!] \in \mathsf{rec}_i(\Gamma)$ by the definition of truth. It follows from the definition of $\mathcal{M}_\varphi$ that there is a α such that (i) $\mathtt{rec}_i\alpha \in \Gamma$ and (ii) $[\![\alpha]\!] = [\![\chi]\!]$. From (ii) we have $\mathcal{M}_\varphi, \Gamma \vDash U(\chi \leftrightarrow \alpha)$ by the definition of truth and hence (iii) $U(\chi \leftrightarrow \alpha) \in \Gamma$ by the truth lemma for CDL plus U. It follows that $\mathtt{rec}_i\chi \in \Gamma$, for otherwise $\neg\mathtt{rec}_i\chi \in \Gamma$ by the maximality of Γ, in which case Γ is inconsistent by (R1) given (i) and (iii). In the other direction, if $\mathtt{rec}_i\chi \in \Gamma$, then $[\![\chi]\!] \in \mathsf{rec}_i(\Gamma)$ by the definition of $\mathcal{M}_\varphi$ and hence $\mathcal{M}_\varphi, \Gamma \vDash \mathtt{rec}_i\chi$ by the definition of truth. The case for $S \preceq_i^{\chi} S'$ is analogous, using (A1) instead of (R1). □

Lemma 3 (Canonicity)**.** *$\mathcal{M}_\varphi$ is a legal testimonial model.*

Proof. $\mathcal{M}_\varphi$ satisfies the first legality condition given axioms (A2) and (A3), the second given (A4), the third given (A5), and the fourth given (R2) and (A6). □

Since the φ with which we began is by assumption not refutable by the DTL axioms, it is contained in one of the maximally consistent sets $\Gamma \in W$. Hence by the Truth Lemma $\mathcal{M}_\varphi, \Gamma \vDash \varphi$, so by the Canonicity Lemma φ is satisfiable in a legal testimonial model, our desired result.

References

L. R. Anderson and C. A. Holt. Information cascades in the laboratory. *The American Economic Review*, 87(5):847–862, 1997.

R. Audi. The place of testimony in the fabric of knowledge and justification. *American Philosophical Quarterly*, 34(4):405–422, 1997.

A. Baltag and S. Smets. A qualitative theory of dynamic belief revision. In G. Bonanno, W. van der Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT7)*, Texts in Logic and Games 3, pages 13–60. Amsterdam University Press, 2008.

J. van Benthem. Dynamic logic of belief revision. *Journal of Applied Non-Classical Logics*, 17(2):129–155, 2007.

S. Bikhchandani, D. Hirshleifer, and I. Welch. A theory of fads, fashion, custom, and cultural change as information cascades. *The Journal of Political Economy*, 100(5):992–1026, 1992.

P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.

O. Board. Dynamic interactive epistemology. *Games and Economic Behavior*, 49: 49–80, 2004.

R. Booth and T. Meyer. Admissible and restrained revision. *Journal of Artificial Intelligence Research*, 26:127–151, 2006.

C. Boutilier. Iterated revision and minimal change of conditional beliefs. *Journal of Philosophical Logic*, 25:263–305, 1996.

J. Cantwell. Resolving conflicting information. *Journal of Logic, Language, and Information*, 7:191–220, 1998.

R. Demolombe. To trust information sources: a proposal for a modal logical framework. In C. Castelfranchi and Y.-H. Tan, editors, *Trust and Deception in Virtual Societies*, pages 111–124. Kluwer Academic Publishers, 2001.

R. Demolombe. Reasoning about trust: a formal logical framework. In C. D. Jensen, S. Poslad, and T. Dimitrakos, editors, *iTrust 2004*, volume 2995 of *Lectures Notes in Computer Science*, pages 291–303. Springer, 2004.

H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Springer, 2008.

C. Geist, B. Löwe, and B. V. Kerkhove. Peer review and knowledge by testimony in mathematics. In B. Löwe and T. Müller, editors, *Philosophy of Mathematics: Sociological Aspects and Mathematical Practice.* College Publications, London, 2010.

A. I. Goldman. Experts: Which ones should you trust? *Philosophy and Phenomenological Research*, 63(1):85–110, 2001.

J. Hardwig. Epistemic dependence. *The Journal of Philosophy*, 82(7):335–349, 1985.

J. Hardwig. The role of trust in knowledge. *The Journal of Philosophy*, 88(12): 693–708, 1991.

C.-J. Liau. Belief, information acquisition, and trust in multi-agent systems—a modal logic formulation. *Artificial Intelligence*, 149:31–60, 2003.

C. List and C. Puppe. Judgement aggregation: a survey. In P. Anand, P. Pattaniak, and C. Puppe, editors, *Oxford Handbook of Rational and Social Choice*. Oxford University Press, 2009.

A. C. Nayak. Iterated belief change based on epistemic entrenchment. *Erkenntnis*, 41:353–390, 1994.

R. Sorensen. Problems with electoral evaluation of expert opinions. *The British Journal for the Philosophy of Science*, 35(1):47–53, 1984.