Disagreement and Meaning: The application of social choice theory to deference, self-designation and vagueness.

Craven, John

Awarding institution:
King's College London

Disagreement and Meaning

The application of social choice theory to deference, self-designation and vagueness

John Craven

Department of Philosophy, King's College, London

June 21, 2023

For my grandchildren

Rosa Eve Ottolenghi

Calder Henry Mather Craven

Vito Benjamin Ottolenghi

Soren Brower Oakes Craven

Franco Emmett Ottolenghi

Thanks

I am very grateful for the patient guidance of Dr Eliot Michaelson who read many drafts of chapters and articles. From him I have learned a lot about the different styles, cultures and thought processes of philosophers and of economists. I hope that he has learned something about this too.

I am also grateful to many members of the Philosophy department at King's College, London who provided a welcome opportunity for a retired university manager to return to academic study through their part-time Masters programme.

My friend and colleague George Burrows, Professor of Music and Theatre at the University of Portsmouth, provided much needed guidance on musical genres, and allowed me to try out ideas concerning disagreement about the use of concepts as a part of the Arts and Humanities Research Network *Musical Theatre and All That Jazz*. The participants tended to agree that there is no common definition of either genre.

Michael Ottolenghi explained the Israeli *Law of Return*, and the complexities of claims to Israeli citizenship.

For many years I have been grateful to many scholars through discussions, seminars and the comments of anonymous referees. These led to the publication of Craven (1992), after which I largely ignored social choice theory because of other commitments. It was a pleasant surprise to reach the conclusion that that approach has something to offer within the philosophy of language.

My family has been hugely supportive, and my wife Laura endlessly patient with the time that I have spent compiling this dissertation and with my distraction when puzzling why proofs go wrong.

Copyright

Abstract

Burge (1979, 1986), Fodor (1994), Greenberg (2014) and others discuss the possibility that an individual's understanding of concepts can develop through deference to the views of others who have greater expertise or experience in their use. This dissertation extends this possibility by examining the individual's response when the influencers do not agree. Their disagreement might result from differences of opinion, from social interaction that involves different influences on the influencers, from different interpretations of evidence or from vagueness which leaves open several possible precisifications that differ in their implied definitions of the extensions of concepts.

In the terminology of Lewis (1969), a language is a convention to which people wish to conform. People can be said to speak the same language, and obey its conventions, even if they do not agree on all uses of concepts. A use of a concept that lies outside the eligibilities allowed by the language is a mistake: our standard example involves concepts that are not natural kinds - namely the use of genre concepts to classify music. There is scope within the language for disagreement about whether Gershwin's *Rhapsody in Blue* [82] is in (the extension of) JAZZ or of CLASSICAL, but locating it in OPERA or HIP HOP does not conform to the conventions of the language.

In studies of ethics, elections, welfare economics and committee decisions, disagreement is often analysed using the techniques of social choice theory. The approach here is to adapt some of these techniques to consider the impact of dis-

agreement on language learning. The tradition of preference aggregation when individuals disagree was formalised by Arrow (1951) and more recently surveyed by Sen (2017). Arrow's celebrated impossibility theorem rules out all but dictatorial forms of social choice. In turn, this led to the formulation of 'domain restrictions' that reduce disagreement to a sufficient extent that dictatorship is avoided, and allow that decisions can be made by following majority views.

A second strand of social choice theory has been developed by Mirkin (1975), Maniquet & Mongin (2016) and others, who examine the consequences of disagreement between individuals about the location of objects in equivalence classes or named categories. We extend this strand of analysis in contexts that involve disagreement about the allocation of objects to the extensions of concepts. We show in chapter 12 that the structure introduced by Maniquet & Mongin is a special case of our own structure.

Our main conclusion (theorems 1 and 2) distinguishes circumstances in which the language restricts the extent of disagreement to the extent that the learner can devise an effective compromise between disparate views (plausibly by following the majority) from circumstances in which the learner must either violate one of the described principles of deference or nominate and follow a single dominant influencer (a dictator).

An alternative route for a language learner to develop her understanding of concepts arises when she considers the properties of the objects under consideration. She might decide, without reference to the views of others, that two compositions are sufficiently similar that she wants to include them in the same genre. This consideration can also involve the need to reconcile diverse evidence. Some properties of *Rhapsody in Blue* support its inclusion in JAZZ, while others support its inclusion in CLASSICAL. A 'dictatorial' conclusion would then entail that the learner allocates music to genre concepts by ignoring all but one of the properties. However, there is a significant difference between learning through the consideration of properties and learning through deference. The former involves an internal consideration by

the learner, who might legitimately recognise that she feels much more strongly about the impact of some properties compared to others. It is much more difficult for the learner to establish and compare the relative intensity with which her several influencers express their views. So 'disagreement' between the implications of properties need not result in an outcome that reflects only one of them.

We examine two other structures within the framework of language. One involves a structure that is often used in discussions of vagueness in which the objects under consideration can be graded in one dimension (the number of hairs on a head, grains of sand …). The focus then is on potential disagreement about the placement of boundaries between the extensions of concepts (BALD and NOT-BALD; BUCKETFUL, HEAP and SAND-DUNE and …). In one sense, vagueness is a possible source of disagreement about where the boundaries are placed. However, the language itself might be vague, allowing the possibility that each individual uses their own private version of the language or idiolect, and that these in turn might be used by the learner to devise a compromise or representative idiolect. At a second stage, the learner devises through deference her own use of the underlying concepts, and these two stages might give contradictory outcomes.

The final structure concerns contexts in which the learner responds to individuals' designations of themselves and others into extensions of the concepts. This might arise with demographic concepts of different genders, races or religions, or in more prosaic examples in which musicians designate themselves or others into jazz-musicians, classical-musicians …. We generalise a result of Kasher & Rubinstein (1997) that applied originally to designations of eligibility for Israeli citizenship. This result does not lead to dictatorship, but still shows that there are circumstances in which there is no scope for compromise between different views because all opinions that conflict with an individual's self-designation must be ignored.

Original contribution

There appears to be little in the literature that examines formally the possible consequences of disagreement in the language, and its impact on those who want to devise their own uses of concepts. Likewise there appears to be little in the literature on social choice that suggests its potential application to the acquisition of concepts in a language.

The social choice structure introduced in the context of disagreement about musical genres is a hybrid of received structures. The conclusions of our theorems 1 and 2 do not appear to have an exact parallel in the received literature. Domain restrictions, such as those entailed by conformity to a language, have not been explored in structures other than that of Arrow.

The literature does not include consideration of social choice when objects are graded, which underlies our theorems 3 and 4, and it does not appear that disagreement which results from the use of different precisifications has been included in discussions of vagueness.

The discussion in chapter 9 involves an extension to multiple classes of the two-class model of Kasher & Rubinstein. It has been published in *Theory and Decision* [17].

Chapter 12 explores domain restrictions in the aggregation of classifications following Mirkin and Maniquet & Mongin and shows that the structure developed in this dissertation is more general than the latter when there are three or more categories. This conclusion has been published in *Global Philosophy* (*Axiomathes*)" [18].

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The question

1.1.1 Concepts

The central theme of this dissertation examines how an individual might develop her understanding of language in the face of conflicting evidence. In particular, it examines how an individual might form and use concepts that are used in thinking and communication as an efficient alternative to listing each instance.

For our purposes, a **concept** is a set of objects that an individual classifies together (or **co-locates**) on the basis of evidence about the similarity or commonality of the objects. Typically within a context, the individuals classify a set of objects into subsets each of which forms a different concept. For the most part, that classification is a partition of the set, but we can allow that an object might be included in more than concept (section 5.1). The evidence includes both the way in which concepts are used by others, and the individual's own assessment of similarities between objects in the set.

For example, in one of the illustrative contexts used here, the evidence that an individual uses entails that JAZZ is a concept that is likely to include Ellington's *Take*

the A-train [81], and CLASSICAL MUSIC is likely to include Vaughan Williams' *The Lark Ascending* [85]. Arguably (section 2.2), the evidence might lead to the inclusion of Gershwin's *Rhapsody in Blue* [82] in JAZZ or in CLASSICAL MUSIC.

The questions addressed here could also be phrased in terms of the extensions of predicates by, for example, co-locating all music for which the individual regards 'x is jazz' to be true. The concept JAZZ is, for our purposes, the same as that extension.

Many of the techniques that are developed here can be applied in circumstances beyond the study of language in which a set of objects is classified into **categories**. The literature (chapter 12) includes examples of the allocation of individuals to tasks and of the classification of a population in a social survey that is based on a variety of attributes. This is reflected in the phrasing of technical results in chapters 10 and 11.

1.1.2 Disagreement

There is potential for an individual to have conflicting evidence about whether a particular object instances a given concept. The evidence might come from other people who influence the individual and who disagree with one another, or from features and properties of the objects that do not, in her view, unanimously support the inclusion of an object within the extension of one concept. The individual then has the challenge of evaluating the conflicting evidence if she is to establish her own view of the extension of the concept.

The fundamental question addressed here is

Is it possible to find a satisfactory compromise when there is disagreement or diverse evidence about extensions of concepts?

Frances & Matheson (2019) remark that

> '[There are] central epistemological issues tied to the recognition of disagreement [that] … philosophers didn't start, as a group, thinking about … in a rigorous and detailed way until the 21st century.

As well as that surveyed by Frances & Matheson, this literature includes Wright (2001, 2006), Kölbel (2004), Belleri (2010), Belleri & Palmira (2013). It has discussed *inter alia* whether people make different statements about the truth of propositions because some of them are right and some are wrong (regardless of whether it is possible to discern right from wrong), or because the disputed statements appear to disagree but are implicitly different statements because they are relative to each individual's own context or tastes. Although this thesis discusses reasons for disagreement in some illustrative contexts, its main concern is different. The fundamental question asks how someone who wants to formulate her own view of concepts copes with the fact that other people do - for whatever reason - use them differently. The individual must comprehend and communicate given the evidence that she has, and that evidence includes contradictory elements. She cannot wait for evidence to emerge (if it ever does) that establishes the truth or falsity of conflicting statements, and she cannot realistically discriminate between different statements by evaluating the personal contexts of those from whom she might learn. She can only hear what they say.

1.2 Contexts

To avoid ambiguity or circumlocution, **Rosa** is the name of the individual whose concept development is under consideration. In any particular context, the set of objects that are considered for inclusion in a concept is **Rosa's set**, and her formation of concepts is a classification of the objects in Rosa's set[1] into the extensions of alternative concepts, given the context. In a particular context, Rosa's set is unlikely to include every possible 'candidate' object. For example, Rosa might be familiar

[1] Usually a partition, but not in section 5.1.

with *Take the A-train*, *The Lark Ascending* and *Rhapsody in Blue* and so she considers all of them when she is devising her musical genre concepts, including JAZZ. However, it is likely that no individual, however experienced, is familiar with everything that might be considered for inclusion in some musical genre, so any individual's understanding of a concept is limited by the extent of their familiarity.

The narrative can involve other individuals, typically those who might influence Rosa in some way who are, unless otherwise specified, Rosa's **influencers**. Given that Rosa is learning from them (or deferring to them) we assume that her influencers are familiar with all the objects in Rosa's set.

Rosa's learning journey begins when she becomes aware that concepts such as JAZZ exist and can refer to some but not all music. This first stage, largely of early childhood development, lies outside the scope of our analysis. We begin at the next stage, when Rosa is likely to show **deference** to influencers who have greater expertise than she does, by using the evidence from them to populate the extensions of the concepts. Burge (1979) and Fodor (1994) examine the role of deference in establishing concepts - but not the impact of disagreement between influencers.

Also, some concepts are used in different contexts, and Rosa can draw parallels between them. If she is considering the extension of TALL, it is important to know whether Rosa is thinking about trees, buildings, people, grass .... She has some carry-over between these contexts because, in every context, a tall object has greater height than a short object. Rosa might initially be familiar with the use of TALL and SHORT for trees, and this carry-over guides her in using these concepts in other contexts, but does not otherwise help her to locate the boundary that separates tall people from short people.

Rosa is likely to hear different statements from her influencers about, for example, which trees are in the extension of TALL or which music is in JAZZ. These different statements could arise because the influencers have different experiences

that lead to different opinions. Some influencers might have seen Californian redwoods, and, even though these are not in Rosa's set of trees, they might have different views about which trees are tall from those held by influencers whose experience does not include redwoods. Chapter 2 documents that critics and commentators disagree about the genre of some musical compositions. This disagreement might arise from different experiences, from different evaluations of the properties required to locate music in some genre, from alternative opinions about the social or demographic origins of the composition, or from the possibility that each of Rosa's influencers is in turn influenced by others. There might be 'schools of thought' within which there is little or no disagreement, but between which there can be (sometimes vehement) disagreement.

At a further stage, Rosa might reduce her reliance on the views of influencers, and determine her view of the genre of a piece of music by considering its **properties**, comparing them to the properties of other music that she has already - perhaps tentatively - categorised. If the properties of two pieces are sufficiently similar, Rosa has good reason to co-locate them in a genre. But 'sufficiently similar' might apply to some properties but not to all, and is often not clearly defined, so that the scope for conflicting evidence remains. The properties of objects do not necessarily all 'point in the same direction'. A painting by, say, van Gogh, might seem to Rosa to be closer to her experience of other IMPRESSIONIST art in its use of shapes, but closer to POST-IMPRESSIONIST in its use of colour. She then faces conflicting evidence.

So Rosa can experience evidence of different opinions of her influencers, or different evaluations on the basis of properties, or both. If she is to decide whether a particular object is to be located in the extension of some concept, she must evaluate these differences, and conflate or aggregate or synthesise their different implications. We suppose that those thought processes lead her to a conclusion, and our ambition is to examine the implications of requiring that the conclusions that she reaches are not inconsistent or capricious. This examination does not look at the thought processes themselves as a psychologist or neurologist might, but

evaluates the **outcomes** of those processes against the evidence available to Rosa, and against actual, potential or counterfactual variations in that evidence.

1.2.1 Language, eligibilities and mistakes

Individuals in a community, including Rosa's influencers, are often said to have a common language, but the potential for disagreement entails that, in practice, they do not all use it in the same way. There is sufficient commonality that members of the community can understand one another, but sufficient diversity that there can be disagreements not only about the substance of what is said, but also about the use of the terms in which it is said.

Chapter 3 discusses how disagreement is constrained by **eligibilities** that are defined by the language of the community. These eligibilities are encapsulated in statements made by people who conform to the language and that can be found in reference manuals such as dictionaries. My walnut tree is eligible for (inclusion in the extension of) TALL and also eligible for NOT-TALL, a knee-pain is eligible for ARTHRITIS and for LUPUS, and *Rhapsody in Blue* is eligible for JAZZ or CLASSICAL. But the conventions of the language entail that a bonsai maple is not eligible to be in TALL, tooth-ache is not eligible for ARTHRITIS and *Rhapsody in Blue* is not eligible for OPERA. The eligibilities arise from the conventional use of terms in the language, and so they limit the extent of disagreement - but not to the extent of enforcing full agreement.

The eligibilities also define **mistakes** in the use of the language. For our purposes, an individual who makes a statement that does not conform to eligibilities is making a mistake within that language. The conventions of the language permit a variety of statements about the classification of an object - and also does not permit others. This is a different definition of 'mistake' from that used in the context of binary logic (see for example, Kölbel, 2004) where it is a mistake to believe or state a proposition that is not true. Eligibilities permit individuals to make apparently different statements when the truth is not known (possibly not knowable), or

where the statements result from different individuals' own unstated contexts. It is reasonable to allow that if the truth of a statement is known, and is not relative to individuals' contexts, then the language ensures that conforming to eligibilities entails that every influencer makes that true statement - and so too will Rosa if she follows that unanimous view. Archetypes of concepts, about which there is full agreement, are discussed section 3.4.

1.2.2 Truth, taste and vagueness

Frances & Matheson (2019) contrast

> disagreements between individuals about belief from disagreements about matters of taste. [Their] focus is on disagreements where there is a fact of the matter, or at least the participants are reasonable in believing that there is such a fact.

The belief that there is, or could be, a 'fact of the matter' might be reasonable, but impractical, in some contexts. For example, several doctors might disagree about whether a patient's joint-pain is caused by arthritis or lupus, but the prevailing state of medical technology does not permit a clear diagnosis. There is a currently unknown correct diagnosis. Long-term developments of medical technology might eventually permit convergence by eliminating all but one possible cause. At the time of the diagnosis, the patient's illness is eligible for ARTHRITIS, LUPUS and perhaps other disease concepts, and the classification of the patient's illness must be based on the evaluation of conflicting evidence in the given state of knowledge.

In other contexts, evidence does not suggest that dialogue or advances in scientific discernment will lead to an agreed outcome. There is no scientific test that leads to unanimous realisation by every critic and commentator that *Rhapsody in Blue* is, or is not, in JAZZ. Continuing disagreement might be regarded as a matter of 'taste' in the sense that each critic's opinion is firmly based on their own evaluation

of conflicting evidence. With one brief exception (section 6.6.3) this dissertation is concerned with the reaction to disagreement (for example by Rosa learning the language from her influencers) rather than with reconciliation between those who disagree or the possibility of a reduction of the extent of their disagreement.

The assertion that there is an un-knowable 'fact of the matter' underlies view that vagueness an epistemic phenomenon (Williamson, 1994). That interpretation argues that it is impossible to adduce evidence to establish the truth value of "that walnut tree is tall" without doubting that there is such a truth value, because there is an un-knowable true position of the boundary between TALL and for NOT-TALL. The walnut tree is eligible for both TALL and for NOT-TALL[2]. Rosa might be faced with several conflicting statements because her influencers use different precisifications of the concepts. The precisifications that her influencers use are an alternative way of expressing the eligibilities of trees for the concepts under discussion. My walnut tree is eligible for TALL and NOT-TALL and there are precisifications that include the tree in TALL, and others that include it in NOT-TALL: it is eligible for both. From Rosa's point of view, the possibility that there is a truth of the matter is irrelevant. If she is deferential, she can only assess her use of TALL from the conflicting evidence of others' statements[3].

For our purposes, multiple eligibilities can arise from vagueness[4], from the limitations of scientific discernment or from potential and permissible differences in taste or opinion. The problem for Rosa is to determine her own view of the extensions of the concepts given the disagreements.

[2]Or maybe for TALL, TALLISH and MIDDLE-HEIGHT but not for SHORT.

[3]Wright 2006 page 40 suggests that a majority view of 'well-qualified judges' might prevail. He rejects the view for reasons connected with his own arguments, but the example of our table 4.1 (page 62) shows that the use of majority views can in any case be inconsistent.

[4]Chapter 8 examines implications of the different possibility that the eligibilities themselves are vague.

1.2.3 Interdependence

In a particular context, Rosa's set can include many objects (musical compositions, trees …) that are, between them, eligible for several concepts. So it is plausible that her locations of two or more objects are interdependent. For example,

- The evidence available to Rosa might unanimously or overwhelmingly indicate to her that two objects are sufficiently similar that they should be co-located in the same concept. This could arise if her influencers signal such similarity, even if they disagree about the concept in which both objects should be located. Or it could arise if Rosa decides that the properties of two objects are sufficiently similar that it makes sense to co-locate them. Or it could arise because the objects (such as two trees in a forest) are indiscriminably different. Similar considerations apply to the separate locations of two objects when her influencers all locate them separately, or when the objects are clearly discriminable. A major consideration is then the need to ensure transitivity: if Rosa considers that an elm and an ash are sufficiently similar in height that they should be co-located, and that the ash and an oak are also sufficiently similar that they should be co-located then she cannot consider that the elm and the oak are sufficiently different that they should not be co-located[5].

- An elm tree might be discriminably of greater height than an ash tree, but within the forest that forms Rosa's set there are trees of much greater height and of much less height than both. The elm and the ash are mid-range within Rosa's set (the forest) and both are eligible for inclusion in TALL and NOT-TALL. But consistency in the use of the concepts entails that Rosa cannot locate the ash in TALL and the elm in NOT-TALL.

[5] A consideration that underlies a sorites paradox in which indiscriminable difference is non-transitive (Dummett, 1975).

- All of Rosa's influencers might distinguish JAZZ from CLASSICAL MUSIC by including some compositions from Rosa's set in each concept, although they do not agree about all of the locations. Rosa might wish to follow this unanimous recognition that neither concept should be empty. If she were to locate compositions sequentially she could be forced to recognise that the last composition in her sequence should be located to avoid an empty concept, or that she must in some other way change her sequence of location decisions.

In each of these circumstances, her location decisions are inter-dependent.

An important consideration in our later analysis is that the possible recognition of inter-dependence between Rosa's locations of the objects is not inconsistent with the possibility that, to keep her thought processes relatively simple, she should locate each object only on the basis of the evidence that she has about the location of that object. There can be inter-dependence in her choice of locations across all of Rosa's set, but independence in the evidence that she uses to make each location decision.

1.3 Social choice

The derivation and evaluation of outcomes that are devised by conflating or synthesising diverse evidence is the central theme of social choice theory. Chapter 12 includes a fuller examination of issues summarised here.

The most familiar applications include the analysis of elections, committee decisions and welfare economics following the work of Arrow (1951). He developed a rigorous approach to disagreement about preferences or personal interests that enriches economic theory by including questions of distribution and welfare alongside the questions of efficiency. In this context, 'efficiency' considers only the extent to which more of one good can be produced within given resources without producing less of another good. This is uncontentious because of the implicit as-

sumption that 'goods' are good. The potentially contentious issue of the distribution of the benefits of efficiency is at the heart of Arrow's examination of social choice.

Arrow's **impossibility theorem** shows that if the method of aggregating the evidence (such as individual preferences) is required to satisfy certain conditions, then no compromise outcome is possible. The outcome must always be the same as the evidence supplied by a single person - a 'dictator'. The conditions are (arguably) reasonable, as is the expectation that there should not be a dictator, but these requirements can be incompatible.

Later contributions, including Wilson (1975) and Rubinstein & Fishburn (1986), show that Arrow's result is one case of more general propositions about the aggregation of potentially diverse evidence. Two other cases are of particular interest here. The structure introduced by Mirkin (1975) considers the aggregation of evidence that is supplied in the form of equivalence relations that define partitions[6] of the set of 'objects' under discussion. The structure in Maniquet & Mongin (2016) concerns the aggregation of evidence that is supplied in the form of classifications of the objects into a number of named categories (such as musical genres). In both cases, arguably reasonable conditions on the formulation of the structure entail that there must be a dictator because the conditions entail that the outcome of the aggregation is determined by only one source of evidence, ignoring all others.

The structures of Arrow, Mirkin and Maniquet & Mongin all involve the aggregation of evidence from several sources about several issues to produce an outcome. In most examples of Arrow's structure, the issues are questions of the form "is object x preferred or indifferent to object y?". If there are n objects, there are n^2 such issues. In Mirkin's structure, the issues are symmetric questions of the form "is object x in the same equivalence class as y?". If there are n objects, there are $n^2/2$ such issues. In Maniquet & Mongin's structure, the issues are questions of the form

[6] By a fundamental theorem of algebra (for example, Lewis, 1965) any equivalence relation that is defined over a set defines a unique partition of that set.

"is object x in category C?". If there are n objects and r categories, there are $n.r$ such issues.

The common features of the three approaches are

The domain. Each source of evidence involves some underlying structure. In Arrow, the evidence from each source on the ranking of the objects (or candidates) must be in the form of a reflexive and transitive binary relation. In Mirkin, the evidence from each source on the partition of the set of objects into equivalence classes must be reflexive, symmetric and transitive. In Maniquet & Mongin the evidence from each source on the classification of the objects must not leave any category empty. The term **unrestricted domain** is used when there is no other restriction on the evidence that can be considered: so, for example, in Arrow, the domain is said to be unrestricted if any individual can state any reflexive and transitive relation. The method of aggregation must define an outcome for all combinations of evidence within the domain.

Unanimity If all the evidence on a particular issue (ranking, co-location of two objects or classification of an object) is the same, then the outcome reflects that unanimous evidence.

Independence The outcome on an issue depends only on the evidence about that issue. In some formulations, independence is supplemented by a monotonicity condition that ensures that if all additional evidence supports an existing outcome, that outcome should not change[7]. By separating the evidence used to derive the outcomes for different issues, independence conditions impose some consistency on the way in which the outcomes for all the issues change when the evidence changes. Independence ensures that the change in outcome is limited to those issues on which the evidence changes.

[7] In other formulations monotonicity is entailed by the other conditions imposed.

So what are often called independence conditions are referred to here as **consistency** conditions.

The range If the domain is otherwise unrestricted, then the outcome is constrained in the same way as the evidence from each source. In Arrow, the outcomes for the issues must be reflexive and transitive; for Mirkin, the outcomes for the issues must be reflexive, symmetric and transitive; for Maniquet & Mongin the outcomes for the issues must not leave any category empty.

The condition on the range of the aggregation process is particularly important in these structures because it provides interdependence between the outcomes for different issues. Consistency conditions entail that the outcome for an issue depends only on the evidence about that issue, but the combination of the outcomes on several issues might not conform to the constraint on the range. For example, the Condorcet (1785) voting paradox indicates a special case of Arrow's result that, with some combinations of individuals' preferences, the outcome of majority voting is not transitive. For each pair of objects taken separately, following the majority yields an outcome, but these outcomes can conflict when three objects (and hence three pairs) are considered together.

There has been a considerable literature, starting with May (1952) and developed by Sen & Pattanaik (1969) that discusses restrictions on the domain in Arrow's structure - that is on the extent of disagreement that can be considered - that prevent intransitivities when majority voting is used. The domain is restricted and dictatorship is avoided. A domain restriction effectively involves a meta-agreement between or an imposed constraint on the individuals (voters) that certain preferences are excluded, and a practical domain restriction indicates reasons why the meta-agreement holds. The most celebrated example is single-peaked preferences (section 12.2.1) which requires that, for example, no voter puts both a left-wing candidate and a right-wing candidate ahead of a centrist candidate in an election. The possibilities for domain restrictions that are sufficient to avoid dictatorship in the

structures proposed by Mirkin and by Maniquet & Mongin are discussed in chapter 12.

Chapter 4 includes a social choice structure that differs from those of Arrow, Mirkin and Maniquet & Mongin[8] and that is arguably more appropriate for examining the development of genre concepts when Rosa is faced with disagreement or conflicting evidence. With an unrestricted domain, reasonable conditions on the way in which Rosa defers to her influencers entail that she must nominate one of them as dominant (a dictator), and so ignore the opinions of all others. The chapter also shows circumstances in which eligibilities that are defined by a conventional language restrict the extent of possible disagreement to a sufficient extent - and hence restrict the domain of Rosa's challenge - so that she can consistently follow the majority view on the location of music into genres.

1.4 Dialect

Social choice theorists have long recognised that their analysis can apply at two levels.

At the individual level, the derivation of an outcome given conflicting evidence can be considered as the challenge to an individual's own thought process. Arrow's structure could be used to derive an individual's ethical position given conflicts between the interests of various participants whose interests matter to her.

At the community level, the derivation of an outcome can be considered as a process within a social group often for practical implication such as, for example, devising an election result or choosing the optimal economic policy within Arrow's structure. Most or all of the contexts suggested by Mirkin and by

[8] The Maniquet & Mongin structure is a sub-case of the structure of chapter 4 (section 12.5).

Maniquet & Mongin involve allocation decisions that are to be implemented (section 12.4.1).

The individual level in the analysis of language is Rosa's challenge of developing her concepts. The community level involves the definition of a **dialect** to describe the characteristic language of a subgroup of a language community who share some characteristic. This might be given by their geographical location, occupational group or participation in some activity in common. Small groups or families might devise their own 'private language' using concepts that are expressed using similar terms to those used by others in the language community but with different extensions.

Even within a subgroup, members might not use concepts in the same way. We can ask for the definition of the representative **dialect** of the group given the diverse usages by its members. Defining the dialect of the group based on its members' usages can have a similar formal social choice structure to the development of Rosa's use of concepts when she takes into account the ways in which others use the language. So the quest for a representative dialect faces similar difficulties when the members of the subgroup disagree.

We can include in the term 'dialect' versions of the language that are used, by agreement, by a subgroup whose existence may be ephemeral. For example, conference delegates debate the interrelation between jazz and musical theatre. The group does not want to be distracted by debate about 'what does JAZZ mean?' or 'what does MUSICAL THEATRE mean?', and so they agree to define the extensions of the genres in a specific way that is as representative as possible of their potentially disparate views. Each delegate accepts this dialect for the purposes of their debate even though it does not necessarily accord in every detail with their own usage. This acceptance of a collective outcome lies at the heart of social systems that are susceptible to analysis using social choice theory. For example, many people accept the outcome of democratic elections even though they personally prefer a losing candidate; people pay their taxes even though they do not think that the tax

system is fair or effective. Conference delegates engage in fruitful discussion even though they differ in what they include in JAZZ.

1.5 Terminology

The discussion so far has identified Rosa, Rosa's set of objects, eligibilities determined as a part of the common language of her community, evidence from her influencers and, in some contexts, evidence of properties of the objects. The outcome that Rosa determines is a classification or categorisation of the objects in Rosa's set, and the classes or categories involved are (within the specified context) the extensions of concepts.

Rosa is using evidence from several sources, and we use the term **profile** to describe a list of evidence. If the evidence changes, then there is a different profile.

It is also helpful to give a name to the relation between a profile of evidence and the outcome that Rosa devises using that evidence: if the profile changes, so also might the outcome. There are many names[9] for this relation within social choice theory which describes the way in which the evidence is synthesised. A somewhat prosaic term is an **aggregator** - because it involves the aggregation of the evidence in a profile. Averaging is a form of aggregation in which a single number is the outcome from a profile of 'evidence' - namely the numbers to be averaged. Rosa is engaged in synthesis or conflation of the evidence, but 'synthesiser' has a very different meaning in electronic music and 'conflator' seems no less prosaic than 'aggregator'. So, in formal terms, we use **aggregator** to describe the relation between the profile of evidence and the outcome from that evidence. At the community level, the aggregator is the relation between a profile of individual idiolects and the dialect of a group.

[9] Social welfare function (Arrow), collective choice rule (Suzumura), aggregation function (Maniquet & Mongin), principle of agreement (Mirkin) …

Arguably 'aggregator' sounds mechanistic - as averaging a set of numbers is mechanistic in that it is easily described. When it is helpful to invoke an easily described form of aggregation, we refer to a **methodology** such as a 'majority methodology' or a 'plurality methodology' in which Rosa's outcome follows the evidence that has the most support. There are many majority methodologies, because Rosa might weight some sources of evidence more highly than other sources (sections 4.6.1, 10.2), or because she needs to find a way of breaking ties when two possible outcomes are supported by evidence of equal weight[10]. The term 'methodology' is not specific about the exact form of the aggregator - an aggregator can be derived from a majority methodology along with whatever tie-breakers or weightings of individual opinions Rosa considers appropriate. And, it is important to note that not every possible aggregator can be based on a simply specified methodology. At its most fundamental, an aggregator consists of a list of the possible profiles that Rosa can consider and, for each profile, the outcome. Two different aggregators associate a different outcome with one or more of the possible profiles. Section 4.2 gives an example in which the number of possible aggregators is very extensive, even though Rosa's set is small and there are few individuals.

Formally, the **principles** that are invoked (conditions to some authors, axioms to others), such as those requiring unanimity, independence, monotonicity or the absence of a dictatorial individual, are conditions or restrictions on the form of the aggregator. Demonstrating an impossibility result involves the conclusion that no aggregator can satisfy all of the principles. It is often easy to devise examples, based on simply specified methodologies, that suggest the impossibility, but proving the result on a case-by-case basis for every possible aggregator is impractical. A major technical contribution of social choice theory is finding a way to avoid this impracticality by providing a direct proof rather than embarking on a full description and evaluation of every possibility.

[10] A committee chair with a casting vote effectively has double weight compared to other members.

1.6 Methodological notes

1. Model-building is a major component of economics, game theory and other theoretically based disciplines, including, of course, aspects of philosophy. For example, Lewis's work on conventions is referenced in chapter 3. Simplifying assumptions are crucial within this tradition, and the applicability of results can always be challenged on the grounds that these assumptions take the analysis far from any plausible real-world problem. All of the formal results here face that challenge - but it is unlikely that any clear conclusion could be drawn from a much more inclusive analysis. Further investigation might moderate some of the specification, perhaps without changing the more challenging outcomes.

2. A discussion that invokes social choice is normally limited in the sense that the only influence on the outcome is the information given by evidence contained in a profile of individual statements. Other information is excluded unless specified otherwise. In particular, the views of Rosa's influencers might be held strongly or weakly, but such statements of intensity are excluded as evidence, unless as in the specific context of chapter 6, they are explicitly included.

3. The questions that are asked here are of the form "Is it possible that …?". In some cases the answer is "no" and then it is necessary to demonstrate that it is not possible to satisfy all the requirements of the question. In other cases, the answer to "Is it possible that …?" is "yes", and this can be justified by devising an example that satisfies all of the requirements. The example might not be the only example that could be used, nor is it in any sense judged here to be the best example, nor is it put forward as the answer that someone seeking to answer the question in practice would use. Usually such examples use an aggregator that is based on a simple methodology, and this might not be the most appropriate aggregator for revealing the subtleties involved in developing concepts, or in communicating when people disagree.

4. One of the approaches taken here is to examine how the outcome of a social choice changes if one or more of the inputs changes. This is, in the jargon of economists, an exercise in comparative statics. It examines how, if at all, the outcome could be different with one profile of evidence compared with another. It is not intended to examine the dynamic process of how a change in the outcome actually occurs.

1.7 Outline

The dissertation uses three illustrative contexts for the challenge of developing concepts using evidence from other members of the community or from evaluations of the properties of objects.

Musical genres. Chapter 2 draws out features of this context that are used as illustrations and that are the basis for the main conclusion of chapter 4. The chapter refers to documented disagreement about the classification of music into CLASSICAL, JAZZ .... Music has properties such as rhythmic structure, tonality, and the use of harmony that may be associated with several different genres.

Spiciness categories. This is presented in chapter 7 as a **graded** context in which Rosa's influencers agree that the dishes (in the sense of components of a meal) can be graded in order from the least to the most spicy but numerical measurement is difficult. The influencers disagree about the location of boundaries between **ranked** spiciness concepts (HOT, MODERATE, MILD ...), and Rosa seeks to develop her understanding of the location of those boundaries. The same context is used in chapter 8 to discuss vagueness, including vagueness of the language.

Demographic categories. Chapter 9 presents this context, in which Rosa uses evidence from other individuals in her community each of whom self-designates and designates others into categories by, for example, race, religion, gender,

> sexuality .... The formal analysis examines the extent to which Rosa might locate some individuals in categories into which they do not self-designate, bringing into focus her possible response to the weight of opinion that opposes some individuals' self-designations.

The first two concepts involve the recognition that the language limits the scope to which individuals can present different evidence. Using examples from the genre context, chapter 3 examines language as an institution that enables the community to function effectively. This draws on the work of Lewis (1969) who contrasts social contracts and conventions as ways of bringing about social cohesion. A social contract, including the legal system in a society, aims to ensure that everyone is motivated to comply for fear of sanctions. A convention, such as the use of money or of language, ensures widespread compliance because it is each person's interest to follow the convention on the basis of being confident that everyone else will comply. In the case of language, compliance includes the use of concepts within the eligibilities that the language establishes.

Chapter 4 formalises Rosa's deference to people more expert than herself to enable her to develop her own use of genre concepts. She does this by following deference principles that relate closely to principles used in much of social choice theory (section 1.3), and this leads to a fundamental conclusion:

- When the eligibilities derived from the language sufficiently restrict the extent of disagreement between her influencers, Rosa is able to compromise between them and satisfy all of the deference principles.
- When the eligibilities permit a greater extent of disagreement, Rosa must in all circumstances follow the opinion of a single **dominant influencer** if she follows the remaining deference principles.

The formal derivation of this fundamental conclusion involves technical analysis, which is detailed in chapter 10.

Chapter 5 explores variations on the theme of chapter 4, allowing that genre concepts might overlap (so that, for example, *Rhapsody in Blue* could be included in both JAZZ and CLASSICAL). Also Rosa might experience new music that she wants to classify into genres in ways that remain consistent with the classifications that she has already made. Finally, section 5.3 reinterprets the conclusions of chapter 4 to examine the definition of a dialect that represents the use of language by a subgroup of the community.

Chapter 6 continues to use the genre context, and examines how Rosa can evaluate potentially conflicting evidence from properties - formalising that, for example, *Rhapsody in Blue* has some properties that she associates with JAZZ, and other properties that she associates with CLASSICAL. This is an internal evaluation by Rosa, which, in its pure form, ignores the opinions of others. Section 6.4 argues that in using this evidence, Rosa can consider her evaluation of the **intensity** with which properties support particular genre locations, and argues that using evidence about intensities is much more problematic when Rosa defers to other people. The chapter concludes with a consideration of the use of both deference to others' opinions and her own evaluation of properties.

Chapter 7 adapts the analysis of chapter 4 to consider the graded context of the spiciness of dishes. Graded objects are allocated to ranked concepts, and the main focus is on the location of the boundaries between the extensions of the concepts. The deference principles are specified in a different way to take account of the difference in the structure of eligibilities defined by the language. If there is agreement that Phaal curry is spicier than Korma, then the eligibilities of these two dishes are interdependent. For example, if Phaal is eligible only for MODERATE, Korma cannot be eligible for HOT. But, as chapter 11 demonstrates formally, the fundamental conclusions of chapter 4 remain.

Discussions of vagueness (including the 'original' sorites formulation) are often conducted in a graded context. At one level, vagueness is another source of difference in the classifications stated by the experts, and this changes little in the anal-

ysis. Chapter 8 raises a more challenging question when it is recognised that the language itself is vague because eligibilities are not precisely defined. A taxonomy of vagueness is developed in section 8.3.2 that is closely related to the distinction in the extent of disagreement that leads to the fundamental conclusion of chapter 7.

Chapter 9 considers the third context. Rosa develops demographic concepts that are based on the self- and other-designations of people in her community. This builds on the work of Kasher and Rubinstein (1997) with a re-interpretation of the features of the structure discussed in section 1.3, together with a principle that prevents at least one form of discrimination between the community members. The conclusion is again that there are circumstances in which Rosa can follow majority or preponderant views, and other circumstances in which she has no alternative to following a liberal principle of locating an individual in the concept in which that individual self-designates.

The analysis has been developed in the three illustrative contexts, but the results can apply more generally within social choice. Chapter 12 relates the results derived here to other contributions in social choice theory, and suggests their use in other applications of the theory.

Each chapter ends with a summary section and chapter 13 presents some overall conclusions.

Chapter 2

Genres in music

2.1 Introduction

This chapter evidences that there is disagreement between expert commentators - critics, performers, musicologists - about genres of music. The evidence quoted focusses on disagreements relating to JAZZ. This is not a treatise on music, and disagreement about other genres is implied rather than evidenced.

The proposition "*Rhapsody in Blue* is in (the extension of) JAZZ" can involve disagreement in three ways:

1. People might be thinking about different compositions or arrangements entitled *Rhapsody in Blue*.

2. People might have different vocabularies. For example, suppose that one person has a fine-grained distinction between genre concepts that, like JAZZ, stem from a particular tradition, whilst the other does not. Then the assertion that a *Rhapsody* is not in JAZZ might have a different basis for each, depending on the alternative genres with which they are familiar.

3. Even if they agree about the identity of the composition and they use the same list of genre concepts, people might disagree about which of their common concepts includes *Rhapsody*.

The chapter addresses these issues in reverse order.

2.2 Disagreement about JAZZ

After the first performance of *Rhapsody in Blue* some critics asserted that it was jazz, and others that it was classical. Gutman (2003) later wrote:

> Even now, with the vantage of retrospect, the *Rhapsody in Blue* eludes convenient classification. Is it classical music with pop elements, or jazz with serious pretensions?'

More generally, there is plenty of evidence of lasting disagreement about what is included within JAZZ. Publications on jazz - and particular those that seek to survey the field rather than focus on a single performer or band - often start with a statement that there are major difficulties in finding anything close to a consensus on what constitutes jazz. For example, Adorno's *On Jazz* (1936) begins:

> The question of what is meant by "jazz" seems to mock the clear-cut definitive answer. Just as the historical origins of the form are disappearing into the fog of the recent past, so its range is disappearing within its ambivalent use at the present moment. For the purpose of providing a crude orientation, one could concede that it is that type of ... music ... [that] is distinguished from what preceded it by its decidedly modern character, a quality which itself, however, is sorely in need of analysis. ... Musically, this "modernity" refers primarily to sound and rhythm, ... Syncopation is its rhythmic principle.

Likewise, Panassié's *The Real Jazz* (1942) begins:

> Almost half a century has passed since the birth of jazz. ... and still the public has not the least idea of what this music really is.

And early in Nanry's *The Jazz Text* (1979) he asserts that

> Among jazz players and listeners great controversy swirls around the sorting out of real jazz, jazz-influence, near-jazz and nonjazz.

This disagreement arises for several reasons.

- There is no scientific test, such as DNA analysis or medical scan, that might serve to reduce or eliminate disagreement as it might when, for example, doctors disagree about the diagnosis of a particular illness.
- Music has many properties, including rhythmic structure, harmony, tonality, instrumentation, scope for improvisation …People can evaluate these properties differently, both in understanding them and in weighting their relative importance.

By locating two compositions in the same genre, an individual is revealing a statement that the two are sufficiently similar that they can be co-located even though (plagiarism apart) the compositions are different. Some 'discipline' is placed on statements of co-location because the ability to classify compositions into genres requires that the binary relation "is sufficiently similar to" is transitive. So asserting that composition a is sufficiently similar to b and that b is sufficiently similar to c entails that a is sufficiently similar to c. If, on reflection an individual does not consider that a is sufficiently similar to c that they can be co-located, he must reconsider his view that a is sufficiently similar to b, or that b is sufficiently similar to c.

In addition to these sources of lasting disagreement about compositions, many discussions of jazz involve its close interaction with social and racial issues, about which commentators disagree. A thumbnail sketch of the development of JAZZ

starts with a musical tradition derived from slaves - and then former slaves and their descendants - in the southern states of the USA. These racial origins remain prominent in the development of jazz and jazz criticism through and beyond the 'Jazz Age' between the two world wars. The evolution of styles is strongly intertwined with the social, economic and technological history of the twentieth century, particularly - perhaps overwhelmingly - in the USA. Racial segregation, prohibition, the Great Depression, World War II, improved transportation, population movements, commercial demands, recording and play-back techniques and radio are all major influences on the continuing story of jazz, and jazz has had a major impact in return. What is, or might be, counted as JAZZ has changed and has been disputed within these social contexts.

2.3 Alternative genres

The jazz tradition is entwined with the development of other genres. In his *History of Jazz* (1998) Gioia remarks

> The ability of African performance arts to transform the European tradition of composition while assimilating some of its elements is perhaps the most striking and powerful evolutionary force in the history of modern music. The genres of music that bear the marks of this influence are legion ... gospel, spirituals, soul, rap, minstrel songs, Broadway musicals, ragtime, jazz, blues, R&B, rock, samba, reggae, funk, salsa, calypso, even some contemporary operatic and symphonic music.
>
> ... blues and jazz have remained intimate bedfellows over the years, ... an intimacy so close that, at times, it is hard to determine where the one ends and the other begins.
>
> Ragtime music rivals the blues in importance ... as a predecessor to early jazz. Indeed, in the early days of New Orleans jazz, the line be-

> tween ragtime and jazz was so fine that the two terms were often used interchangeably.

This reflection adds to the possibilities for disagreement about the extension of JAZZ, and also points to the second source of disagreement given in section 2.1. Disagreement can arise because some people have richer vocabularies that recognise more genres than others. A simplifying assumption (cf section 1.4) of our analysis removes this reason why people might state different views: **we assume that every individual involved uses the same list of genres**. So when Rosa bases her own classification of compositions into genres on the statements of others, she uses her own list of genres, and interprets others' statements within this list. She might later develop a richer vocabulary, and then she sets up a different challenge, but that goes beyond the scope considered here.

2.4 Originalism

For the identification of most compositions, it seems straightforward to refer to the original (or some canonical) performance. This identifies *Rhapsody in Blue* as the composition by George Gershwin, named by his brother Ira, that was first performed in the Aeolian Hall, New York, on February 12, 1924. This route to agreement about the concept RHAPSODY IN BLUE is essentially originalist (Sainsbury and Tye, 2011, 2012) because individuals construct the concept by reference to some original identification, either directly or indirectly.

This recourse to originalism in identifying musical works (or the concepts that their names identify) is arguably a feature of classical music more than of jazz or other forms that derive from the traditional music of particular ethnic, racial or geographical groups. Some jazz is adapted, arranged and 'borrowed' in ways that depart some distance from the original composition or performance. There is likely to be a lack of clarity about the boundary between those performances that can be encompassed under the original title, and those that should be regarded as new

compositions in their own right, and perhaps in a different genre from the original. *Some Day my Prince will Come* [87] is now regarded as a jazz standard but plausibly was not when it was originally introduced in *Snow White and the Seven Dwarfs* because musical film/theatre is arguably a genre in its own right. Clearer boundaries arguably arise for Rachmaninov's *Prelude in C♯ Minor* [89] which is likely to be regarded as classical in performances by the composer, and as jazz in the King Cole arrangement [84]. The Modern Jazz Quartet's arrangement of J.S.Bach's *Air on a G-string* [86] is very different from the classical performance by Anne-Sophie Mutter [77]. Substantially different arrangements are different compositions for the purposes of illustration here.

2.5 Historic origins of genre concepts

Originalism can help to justify that there is agreement about the identity of compositions. It is much less plausible to use some single originating 'naming event', by an individual or in a definitive text to give meaning to a genre concept such as JAZZ[11]. Someone, probably in or observing the black community in the southern US around 1900, first used the term 'jazz' in commenting on some performance[12]. But it is difficult to see how that original user could have had dispositions that would enable him or her to classify all of the music that has subsequently been devised and so guide anyone to a modern-day usage of the concept. The universe of music changes all the time. New works may be classed in JAZZ, and some might be reclassified into other genres. To remain originalist, we would need to find a clear trail from this necessarily very limited original use of the term to its use at a later date in

[11]The legal equivalent allows a "collective" as the originator. Originalists assert that the US Constitution should be interpreted as the Founding Fathers intended. Greenberg (2011) gives a comprehensive discussion of parallels between legal and language questions, including the recognition that it is not straightforward to discern legislators' intentions. The founding fathers' statement might have been vague, or used terms whose conventional usages have subtly changed, or partial in the sense that the world has changed in ways that they could not consider. There are divergent opinions amongst lawyers.

[12]It might be a corruption of some earlier term, which has its own ancestry. Or perhaps it was Jelly Roll Morton who claimed (implausibly) to have invented jazz [36].

a much broader context. It seems highly unlikely that people agree on the nature of this trail because, unlike the meaning of Rhapsody in Blue or Beethoven's 9th Symphony [78], the meaning of jazz changes as people discover music that is new to them or to the world.

Some commentators, for example, Panassié (1942) and Goffin (1946), argue that a performance can be classed as 'authentic jazz' only if there is a robust connection, or sequence of connections, from the music of slaves and emancipated slaves in the United States through to that performance. Also, views have been expressed that a necessary condition for a performance to be classed as jazz is that the performers are black. The connections between race, identity and music are complex, and not limited to the high-profile context of racial segregation in the United States. Is Klezmer less authentic if it is not played by a Jewish band? These opinions do not trace the concept back to its single originator, but to the culture from which it arose. However, music derived from that of slave communities is now classed in several genres as the comments by Gioia quoted above indicate.

It is difficult to see how the originator of the term 'jazz' could be a relevant influence (even less, the sole influencer) on a present-day understanding of jazz. The universe of music grows, so that the original conception of jazz and other genres is unlikely to encompass all that might reasonably be included in it. Further, there is likely to be disagreement about the identification of connections back to that original user or to the culture where jazz began. The challenge of reconciling disagreement about the genre of a composition is not likely to be met simply by reference to a single originator, or by agreement about the trail from the origin to the present.

2.6 Properties of compositions

Chapter 6 discusses ways in which Rosa might structure her thinking if she evaluates the properties of compositions, and how those properties suggest, perhaps not unanimously, the genre concepts that she uses.

The musical properties usually associated with JAZZ include:

syncopation: normally off-beat notes are emphasised;

improvisation and variation: significant parts of the melody and harmony are left to the discretion of performers to produce their variations on a basic theme;

blue notes: notes are flattened (usually) from their conventional pitches, arising in part from a compromise between traditional classical seven-note and African five-note scales;

multiple rhythms that run simultaneously;

swing: which in this definition (as opposed to its possible use as the name of a genre involving big bands) refers to the co-ordination needed to ensure that multiple rhythms, improvisation and variations produce a coherent sound rather than a cacophony.

instrumentation: often including saxophones and brass, usually bass, often other wind instruments, not unusually piano and voice …

For the purposes of chapter 6, we assume that it is possible to identify properties of other genres as well. These are likely to overlap with the properties associated with JAZZ. To avoid musicological diversions that are not the focus of this dissertation, we assume that there is a finite list of properties of all compositions under discussion. Chapter 6 considers how the properties of a composition relates to those that are associated with each genre and so affects Rosa's concepts.

2.7 Summary

The purpose of this chapter has been to introduce an illustrative context in which there is strong evidence that even experts, who include Rosa's influencers, disagree. There are several possible reasons for this that might be summarised as 'different opinions' - but that can result from the response of an influencer to those who have influenced them, or from different evaluations of the properties of compositions, or from different interpretations of the historical or social context of the compositions.

Any such context is likely to be simplified or stylised, and its purpose is to replace statements such as 'the learner locates object x in concept C' with 'Rosa locates *Rhapsody in Blue* in JAZZ'. The remainder of this dissertation could be phrased in the more abstract and austere way at the cost of many circumlocutions and much potential for confusions.

Chapter 3

Language

3.1 Social institutions

Despite the challenges that arise from diversity and disagreement about the use of terms, we still say that a community speaks a common language. The use of a common language is a social institution that is a major contributor to the ability of members of a community to benefit from interaction. Arguably it is the most important contributor, because the co-ordination that arises through other institutions such as mutual respect for rights, the use of money, a legal framework, a political system and the division of labour cannot be achieved without communication.

Lewis (1969) introduces two structures that can give rise to conformity of behaviour between members of a community. The first is a **convention** which Hume (1748) identifies as:

> a sense of common interest; which sense each man feels in his own breast, which he remarks in his fellows, and which carries him, in concurrence with others into a general plan or system of actions, which tends to public utility.

Lewis models this in a game-theoretical framework in which every member of the community has two options: to conform or not to conform. A co-operative equilibrium occurs when everyone prefers to conform on the condition that everyone else conforms as well. If everyone conforms, no-one wants to change her chosen option. Widespread non-conformity leads to an outcome for almost every member of the community that is much worse than the outcome from near universal conformity, and arguably widespread non-conformity is contagious, as each individual sees that others are gaining through their actions.

The use of money is largely the result of the convention that each person uses money because she believes confidently that others will accept it too, and this confidence is reinforced because people are generally aware that others in turn have a similar belief. There are some laws governing the use of "legal tender" and money more widely defined (cheques, debit cards, bank transfers), but these are generally designed to protect the state's monopoly on profitable money-creation through, for example, minting coins and notes, approving banking regulations and quantitative easing. The potential for penalties can deter people from writing cheques when their bank balance is insufficient to meet their obligation, and banks from creating too much credit. People do not think much about those things when deciding to use money in routine transactions because they trust largely in the convention that the money that they receive will enable them to engage in further routine transactions because others will accept their money. Some people may not use cheques or bank transfers, perhaps because they do not trust them to the extent that they trust the use of bank-notes, but conformity is widespread and rests largely on the recognition of mutual conformity to a convention. If people lose confidence in the use of money because of hyper-inflation or because the money in question is a failing cryptocurrency, then the decline in the use of that form of money can rapidly be contagious.

The second structure considered by Lewis is a **social contract** between the governing authority and members of the community. Hobbes' social contract (1651)

between the authority and community members involves laws in a justice system that, at least minimally, protect persons and property. The alternative is a state of nature in which people are not protected from violence and theft. In return for the legal protection, the authority extracts surplus, but not beyond the point where community members regard the state of nature as a better option than very high taxation. A democratic social contract allows citizens to choose the authority, and potentially restrain its extraction of surplus. So most people do not threaten others' persons or property: they understand the advantages of conforming to the behaviours that enables much of personal and economic life to continue successfully, and are deterred by the penalties for non-conformity.

3.2 Conventional language

The use of language is a social institution that is largely unregulated through the justice system. It is conventional and is not the subject of a general social contract. Meanings recorded in dictionaries or given by experts are not imposed by law. Of course, the laws that protect persons and property and regulate transactions are phrased in the language, and some of the actions thus proscribed might be verbal. For example, laws attempt to protect people from defamation and from the release of their medical data to unauthorised others. So there are true sentences that cannot legally be uttered in some circumstances. Some might be deterred from using some terms through social disapproval, as, for example, L'Académie Française might deprecate the use of 'email' instead of 'courriel', but not with legal authority. In general, people conform in their use of language because it is beneficial to do so, given that others also conform, and not because they are motivated by fear of penalties.

The benefits of a co-operative equilibrium might not be achievable by language users if communication between them is significantly constrained. A prisoner's dilemma illustrates this: both captives would prefer the co-operative outcome but cannot achieve it because they are held separately. Despots try to limit the use

of social media and mobile communication when carrying through a coup d'état. Occupying forces might ban the use of a traditional language. However, it is difficult to find examples of situations in which the legal system has sufficient power to prevent the use of a common language without restricting channels of communication. Perhaps the only narrative in which this happens involves divine power. *Genesis* 11 describes a power-play by God to restrain the ability of the people to act in unconstrained concert:

> And the LORD said, Behold, the people …have all one language; …and now nothing will be restrained from them …let us go down, and there confound their language, that they may not understand one another's speech.

Their disapproved behaviour was prevented because they were deprived of their common language, not because channels of communication were banned. They could talk to one another but without mutual comprehension.

Compared with the use of money, language use is complex because each person experiences a significant life-long learning process, and there are likely to be many dialects in which individuals participate temporarily. But in general it is beneficial to conform most of the time to the language of the community in the expectation that sufficient others have learned to do so as well, and that yet more will learn in the future.

3.3 Eligibility

Received discussions of conventional use of language (surveyed in Rescorla, 2019, for example) often do not examine the possibility of disagreement within the language, particularly between 'experts' who conform to its conventions. But it hardly seems fruitful to argue that we speak different languages if I say "*Rhapsody in Blue* is jazz" and you say "*Rhapsody in Blue* is classical". Each of us can understand the

other's claim, each can acknowledge the right of the other to a contrary opinion. We disagree within the language.

However, if a rational individual who is very familiar with *Rhapsody in Blue* were to persist in the statement that it is an opera, we would conclude that they are speaking a different language, in which OPERA includes compositions that we (speaking our language) would certainly exclude. The two languages are distinguished once the possibilities that mistakes stem from ignorance and misunderstanding are removed. The languages differ in the **eligibilities** that permit the inclusion of each composition in some but potentially not all genres.

Some aspects of language are incorporated into dictionaries, and those who conform to the language do not use it in ways that conflict with "dictionary definitions". But they need to learn to conform, given that people use many terms without first consulting a dictionary. In *Individualism and the Mental*, Burge (1979) gives an example:

> Generally competent in English, rational and intelligent, the patient reports to his doctor his fear that his arthritis has now lodged in his thigh. The doctor replies by telling him that this cannot be so, since arthritis is specifically an inflammation of joints. Any dictionary could have told him the same. The patient is surprised, but relinquishes his view. …

Burge has previously asserted that the patient has been using ARTHRITIS in ways that conform to dictionary definitions, so the concept is not new to the patient. But the patient then makes a mistake in using the concept when a new situation arises. By 'relinquishing his view', the patient learns to conform with the convention of the language that thigh-pain is not eligible for ARTHRITIS.

In this example, every competent doctor would make the same correction, but might disagree within the eligibilities of the language about the cause of the thigh-pain, perhaps diagnosing muscle-damage or deep-vein thrombosis. Only people

who persist in the face of expert opinion in attributing the thigh-pain to arthritis can be said to be speaking a different language.

Dictionary definitions (taken here from the *Cambridge Dictionary* online [10]) of musical genres include

> **jazz** ≡ a type of modern music originally developed by African-Americans, with a rhythm in which the strong notes often come before the beat. Jazz is usually improvised (= invented as it is played).
>
> **classical music** ≡ music that is considered to be part of a long, formal tradition and to have lasting value (*sic*).
>
> **opera** ≡ a musical play in which most of the words are sung, or plays and music of this type.
>
> **oratorio** ≡ a piece of music for orchestra and singers that tells a story, usually on a religious subject, without acting.

Other dictionaries give alternative definitions, but without contradicting the propositions such as:

> "*Rhapsody in Blue* is eligible for JAZZ or for CLASSICAL but not for OPERA or ORATORIO"
> Handel's *Semele* [83] is eligible for OPERA and ORATORIO but not for JAZZ"

In summary

- A language includes a collection of eligibilities. Each composition is eligible for one or more genre.
- An individual conforms when she includes a composition in a genre for which it is eligible.
- Disagreement occurs without non-conformity when two people include a composition in different genres for which it is eligible.
- The language defines the scope for disagreement.
- A mistake occurs when an individual includes a composition in a genre for which it is not eligible.
- Every individual recognises the advantages of conformity, and so is willing to learn to avoid mistakes.

A different language can use the same set of concepts, but with different eligibilities. Burge extends the arthritis example to consider the difference between two people in different language communities (worlds) who have identical experiences up to the moment when they visit their doctor. Rosa_1 inhabits a world where the language is $\mathcal{L}_1$, in which thigh-pain is not eligible for $\textsc{arthritis}_1$; Rosa_2 inhabits a world that has language identical to $\mathcal{L}_1$ except that its language $\mathcal{L}_2$ allows that thigh-pain is eligible for $\textsc{arthritis}_2$. By saying that "my thigh-pain is possibly caused by arthritis" Rosa_1 makes a mistake in $\mathcal{L}_1$, and her thoughts about arthritis will be corrected by referring her to the dictionary. Rosa_2 does not make a mistake in $\mathcal{L}_2$, and can continue without correction.

The purpose of Burge's comparative worlds and languages is to show that our correct use of concepts depends in part on features (the language of our community in this case) that are external to ourselves. That is not the topic here, and later chapters show that both introspection about the properties of compositions and external influence through deference can have roles in determining how Rosa devises

her concepts. Deference to those who do conform ensures that Rosa conforms too, whereas reliance on introspection alone might not do so.

3.4 Communication and archetypes

Considerations of communication between individuals within a language raise other challenges.

As a simple example, suppose that Rosa is familiar only with compositions a, b, c, d all of which are eligible for both JAZZ and for CLASSICAL. Rosa locates $a\&b$ in CLASSICAL and $c\&d$ in JAZZ. Her friend Vito locates $a\&b$ in JAZZ and $c\&d$ in CLASSICAL. When they communicate, is the diversity of their statements the result of disagreement about the allocation of music to genres, or does it arise because they agree about everything except the labels that they attach to the genres? On what basis can Rosa and Vito communicate about the genres?

Amongst aficionados of jazz, there is significant agreement about **standards** of JAZZ - music that everyone regards as jazz. Ellington's *Take the A-train* [81] is such a composition. More generally we can identify **archetypes** of other genres: we can say that Vaughan Williams' *The Lark Ascending* [85] is an archetype of CLASSICAL and Britten's *Peter Grimes* [80] is an archetype of OPERA. In general:

> **in a given language, a composition is an archetype just if the language entails that it is eligible for exactly one genre.**

If Rosa and Vito had been familiar with these archetypes, and conform to the language in using them, then they both define JAZZ as the genre that includes *A-train* and CLASSICAL as the genre that includes *Lark*. There is a common basis to their use of concepts that gives them grounds for communication, even if Rosa thinks that c is co-located with *A-train* and Vito thinks that c, is co-located with *Lark*. It is then clear that they disagree about the allocation of c to commonly defined genres, rather than about the labels that they attach to sets of compositions.

In some contexts, archetypes are identified by, for example, expert botanists (Fodor, 1994) who all agree that x is an elm tree, perhaps on the basis of some uncontroversial scientific test. Rosa can identify and defer to these experts and so recognise the archetype elm. In other contexts, there are no professionally qualified experts and Burge (1986) writes

> Meaning-giving characterizations, for ordinary terms [such as sofas], are usually arrived at through reflection on archetypical applications. …To provide meaning, a proposed normative characterization must accord with archetypical applications and must treat the characterizations that competent users actually give, as at least approximations to the norm. (pages 703-04)

This asserts that, in these cases, communication between "competent users" can establish archetypes, and Rosa can learn about them as she participates in her language community. Archetypes serve to ensure that there is some commonality between users of a language.

3.5 Summary

The community in which Rosa learns about concepts has a language which is the product of a convention to which community members wish to conform, provided that sufficient others conform, to enable communication. Language is a vital social institution which is unlikely to arise from a motivation to avoid the sanctions imposed by the authority of the community.

Conformity to the language does not eliminate the possibility of disagreement, but restricts it according to the eligibilities of each composition for inclusion in each genre. Anyone who uses a concept in a way that does not conform to the eligibilities entailed by the language makes a mistake, and can be corrected if they defer to the usage of those who do conform. There can be no disagreement by those who

conform to the language about the location of compositions that are archetypes of genre concepts.

Chapter 4

Deference

4.1 Introduction

Early in her journey of discovery about the meaning of a concept, Rosa is unlikely to have sufficient experience or confidence to think clearly about meanings without guidance from others, or to be willing to enter conversations that require that she has a reasonably mature opinion.

Burge's (1979) example (page 54) illustrates that deference to an expert (the doctor) or reference to dictionaries can correct mistakes and enable an individual to conform to the language. We can suppose that an expert influencer conforms to the language, and so he does not locate a composition in a genre for which it is not eligible. However, this does not fully resolve matters when there is scope for disagreement within the eligibilities of the language. No influencer supports locating *Rhapsody in Blue* in OPERA or Rosa can conclude that the dictionary definition of OPERA ("a musical play in which most of the words are sung") does not apply to *Rhapsody*. But if there is disagreement between her influencers, she cannot learn immediately whether she should include *Rhapsody* in JAZZ or CLASSICAL.

This chapter questions how Rosa can devise her own opinion of the meaning of concepts only by deferring to the views of influencers whose opinions she derives from their statements or writings. In some contexts, influencers might be 'badged' through qualification as medical practitioners, professional scientists or in other ways. In the context of musical genres, Rosa defers to a group of people because of their roles as authors, commentators, critics, musicologists, performers and so on. For our purpose, this influence is one-way: no influencer uses any information from Rosa in his own definition of concepts. The only inputs that Rosa makes to the process are to nominate her influencers, to list the compositions under consideration (Rosa's set) and to establish a list of genres (section 1.2).

Although Burge, Fodor and others consider deference in contexts that suit their purposes, deference and disagreement have not be considered together in the literature. Indeed Greenberg (2014, page 160) remarks that

> Philosophical appeals to deference or community are typically vague and underdeveloped.

The aim here is to provide a rigorous treatment of deference when there is, or can be, disagreement.

4.2 Some examples

Rosa's task in responding to the views of her influencers need not be difficult if her concern is only to decide on the compositions to include in JAZZ, with the implication that all others are consigned to NOT-JAZZ. She can then follow the majority opinion, choosing one influencer as a tie-breaker if needed, and possibly weighting experts to allow that she considers some to be more influential than others. This gives her a straightforward method, and she can apply it in turn to each composition that she considers. The aggregator is responsive to every influencer because each can be a 'marginal voter' in the sense that there are circumstances in which Rosa will change her response if that influencer alone changes his opinion.

In table 4.1, *Rhapsody in Blue* is eligible for JAZZ and for CLASSICAL and Lee Morgan's *The Sidewinder* [88] is eligible for JAZZ, BLUES and SOUL.

influencers	*Rhapsody*	*Sidewinder*
1,2,3	CLASSICAL	JAZZ
4,5	JAZZ	BLUES
6,7	JAZZ	SOUL
majority	JAZZ	no genre
preponderance	JAZZ	JAZZ

Table 4.1: Majority and preponderance classification

JAZZ is clearly the majority location for *Rhapsody*, but there is no majority location for *Sidewinder* There is a majority for each of NOT-JAZZ, NOT-BLUES and NOT-SOUL as locations for *Sidewinder*, and so following these majority opinions entails that Rosa does not locate *Sidewinder* in any genre.

If Rosa locates a composition in line with the best supported view of its location, using an aggregator based on a preponderance[13] methodology, there is a possibility that she co-locates two performances in a genre even though every influencer regards them as sufficiently dissimilar that they are not in the same genre. In the example of table 4.1, more influencers locate *Sidewinder* in JAZZ than in either of the other genres for which it is eligible, and more experts locate *Rhapsody* in JAZZ than in CLASSICAL. So Rosa reaches the conclusion that the compositions should be co-located in JAZZ, even though no influencer states that they are sufficiently similar to co-locate them in any genre. If Rosa wants to devise a defensible classification that is based entirely on deference to these influencers, she cannot quote any influencer in support of the view that *Rhapsody* and *Sidewinder* are sufficiently similar that they can be co-located.

[13]The standard usage in voting theory is: majority ≡ more than 50% of the voters/influencers; plurality ≡ largest group and ≤ 50%; preponderance ≡ majority or plurality.

A third easy-to-describe possibility is that Rosa's aggregator entails that she adopts the modal classification, which is the classification of all the compositions that is stated by the largest number of influencers, with tie-breakers where necessary. In table 4.2, Rosa's set consists of Morgan's *The Sidewinder* [88], *Rhapsody in Blue* and Bernstein's *Prelude, Fugue and Riffs* [79]. *Sidewinder* is eligible for JAZZ, BLUES and SOUL, *Rhapsody* and *Prelude* are both eligible for JAZZ and CLASSICAL.

influ-	4.2a			4.2b		
encer	*Sidewinder*	*Rhapsody*	*Prelude*	*Sidewinder*	*Rhapsody*	*Prelude*
1,2	JAZZ	CLASSICAL	JAZZ	JAZZ	CLASSICAL	JAZZ
3	BLUES	CLASSICAL	JAZZ	BLUES	CLASSICAL	JAZZ
4	SOUL	JAZZ	CLASSICAL	SOUL	JAZZ	JAZZ
5	SOUL	JAZZ	CLASSICAL	SOUL	JAZZ	CLASSICAL
mode	SOUL	JAZZ	CLASSICAL	JAZZ	CLASSICAL	JAZZ
	Influencer 5 is tie-breaker			No tie-breaker needed		

Table 4.2: Modal classification

The modal classification for the profile of column 4.2a is that of influencers 4 and 5 because Rosa has nominated 5 as her tie-breaker[14]. The modal classification for the profile of column 4.2b is that of influencers 1 and 2. The only difference between columns 4.2a and 4.2b is that influencer 4 changes his location of *Prelude* from CLASSICAL to JAZZ. This leads Rosa to adopt a classification for profile 4.2b in which every composition is in a different genre compared to example 4.2a. A single genre change of one composition by one influencer leads to multiple genre changes by Rosa, and, arguably, to instability or discontinuity in Rosa's use of concepts. She cannot locate one composition without evidence concerning all other compositions. For example, her evidence about the location of *Rhapsody*

[14]Note that the modal aggregator with profile 4.2a leads Rosa to locate *Prelude* in CLASSICAL and *Rhapsody* in JAZZ even though a majority of influencers disagree. She adopts the classification of 4&5 because 1,2&3 do not all agree about the genre of *Sidewinder*, and she has nominated 5 as her tie-breaker, so that 4&5 'out-vote' 1&2.

is unchanged between the two profiles, but her view of the location of *Rhapsody* is different. This interdependence in the methodology is likely to be very burdensome when Rosa's set contains a large number of compositions. In social choice terms, Rosa's aggregation methodology fails the independence or consistency requirement of section 4.3.3 below.

These are examples of ways in which Rosa might use the statements made by her influencers to derive her own classification. Each such aggregator is open to potential criticism: following the majority can leave some compositions allocated to no genre; following the preponderance can lead Rosa to make statements of similarity between compositions that are not supported by any influencer; following the mode requires that Rosa knows every part of the profile of her influencers' classifications before she can locate any composition in a genre. The examples are easily described, but there are potentially many other ways in which Rosa could respond to her influencers' classifications, some of which might avoid the sorts of difficulties raised by the examples.

An exhaustive investigation is impractical. Even with the limited example of table 4.1 (2 compositions, 4 genres, 7 influencers) each influencer could state one of six possible combinations of genre locations that conform to the eligibilities of *Sidewinder* and *Rhapsody*. So the influencers could together express 6^7 distinct profiles. Before discovering the influencers' classifications, Rosa might not conform to the eligibilities, so she could place each composition in any one of 4 genres (CLASSICAL, JAZZ, BLUES & SOUL), giving her 2^4 possible classifications. An aggregator then consists of a list of each of the 6^7 profiles each associated with one of the 2^4 classifications. This gives her $6^7 \times 2^4$ (= *c*4.5 million) distinct aggregators in total. In some of these (including any in which she locates a composition in a genre for which it not eligible) Rosa does not reflect some unanimous view of her influencers. Others, like the modal methodology, have major, and arguably unrealistic, information requirements. The challenge is to establish which, if any, of these aggregators is immune to these sorts of potential criticisms. Avoiding an exhaustive examina-

tion of millions of possibilities requires the specification of some principles that can be used to judge the way in which an aggregator maps profiles to outcomes.

4.3 Deference principles

This section sets out some plausible deference principles: they are presented more formally, and their implications are derived in chapter 10. These principles are interpretations of the 'standard' conditions used in social choice: ***unanimity*** and ***independence*** (section 1.3).

4.3.1 Class unanimity

The main challenge is to examine the implications for Rosa's development of concepts when her influencers disagree. But any comprehensive analysis needs to include her reaction when there are aspects of the challenge about which the influencers do not disagree.

> ***Class unanimity***: If no influencer locates composition x in genre G then Rosa does not locate x in G.

Class unanimity trivially entails that Rosa locates x in G if every influencer locates x in G, because no influencer locates x in any other genre.

By assumption, Rosa's influencers conform to the language so that ***class unanimity*** ensures that Rosa does not locate a composition in a genre for which it is not eligible, and so prevents her from making mistakes. However, Rosa's information base is only that contained in the profile of influencers' classifications. So she cannot distinguish whether their unanimity arises from their conformity to the language in which a composition is eligible only for one genre, or whether it is coincidental and there are eligibilities that allow them to disagree about the location of that composition, but they do not actually so so.

4.3.2 Pair unanimity

If Rosa decides to move beyond deference as she becomes more experienced (chapter 6), she is likely to take into account the possibility that two compositions are sufficiently similar that she co-locates them in a genre. When she defers to influencers, she can gain evidence to support co-location, or to support the separate location of two compositions. The minimal support for 'sufficient similarity' is that some influencer co-locates the pair, and the maximal support is that no influencer locates the pair separately. Maximal support occurs even if different influencers co-locate the two compositions in different genres. Then Rosa follows

> ***Pair unanimity***: (i) If no influencer co-locates x and y then Rosa does not co-locate x and y;
>
> (ii) If every influencer co-locates $x\&y$ then Rosa co-locates $x\&y$.

4.3.3 Class consistency

The discussion of the modal methodology exemplified in table 4.2 centred on the view that it was unreasonable to expect that Rosa would use her influencers' opinions about all of the compositions in Rosa's set to determine her location of one of them. It is much more plausible to assume that she develops her concepts by separating her location decisions so that she can locate, say, *Sidewinder* as long as she knows her influencers' genre locations of *Sidewinder*.

This piecemeal approach greatly simplifies Rosa's response when some influencer changes his classification. An influencer might state a different genre location for *Rhapsody* because he re-evaluates the features of the composition, or because he is himself influenced by others. Or he might become familiar with some other work (not in Rosa's set) that changes his perception of what JAZZ includes, and this in turn leads him to relocate *Rhapsody* from CLASSICAL to JAZZ or *vice versa*. Then

Rosa faces a profile that differs only in that an influencer has relocated *Rhapsody*, and she wants to react to it, preferably in an efficient and consistent way.

This argument supports Rosa's use of:

> ***Class consistency***: If there is no change in any influencer's stated location of x then Rosa does not change her location of x.

In the terminology of social choice this is an independence condition: Rosa's location of x is independent of every influencer's location of any other composition. She relocates x only if some influencer relocates x. As well as limiting her response to the statements of her influencers, this deference principle entails that Rosa does not respond to changes in other information (such as her own reflection on the properties of two compositions) to form her locations.

Suppose that some influencer relocates *Prelude, Fugue and Riffs* from CLASSICAL into JAZZ, but does not change other locations. This might entail that, after the change, every influencer co-locates *Prelude* and *Rhapsody in Blue*. Then if Rosa uses ***pair unanimity*** she co-locates the two compositions after the change. ***Class consistency*** entails that she achieves this (if it is not already her conclusion before the change) by re-locating *Prelude* to co-locate it with her previous location of *Rhapsody* and not by re-locating *Rhapsody*.

4.4 Connected language and linked genres

To avoid a context that effectively consists of two or more separate agendas, we assume that the language entails eligibilities that in turn **link** all pairs of genres within Rosa's set of compositions. This excludes the possibility that, for example, Rosa's set consists of some compositions that are eligible only for JAZZ, SOUL and BLUES, and the remaining compositions that are all eligible only for OPERA and ORATORIO.

The eligibilities and the extent of Rosa's set are derived from the language that she and her influencers use. The formal requirement for the language to be **connected** within Rosa's set is

- Two genres are **directly linked** if there are two compositions in Rosa's set that are independently eligible for both genres.
- Two genres are **indirectly linked** if there is a sequence of pairs of genres from one to the other that are directly linked in Rosa's set.
- The language is **connected** within Rosa's set if every pair of genres is directly or indirectly linked.

In this definition, x and y are independently eligible for two genres if the location of x is not contingent on the location of y, and *vice versa*. Some counter-examples (section 10.7) to the results proved here can arise if exactly one composition is eligible for two genres. Then the language is not connected according to the above definition. This special circumstance requires that two genres are similar enough that some composition is eligible for both, but Rosa's set is restricted enough that it excludes all but one such composition. Others might exist in the universe of compositions.

The possibility of indirect linkage allows that, for example, Rosa's set need not include compositions that are both eligible for CLASSICAL and POP provided that there are two compositions in Rosa's set that are independently eligible for CLASSICAL & JAZZ, two that are independently eligible for JAZZ & ROCK and two that are independently eligible for ROCK & POP.

4.5 Two contrasting theorems

Section 1.3 heralded the possibility that the extent of possible disagreement affects whether Rosa can find an aggregator (that enables her to incorporate the views of all of her influencers, or whether she must nominate a dominant (dictatorial)

influencer so that her methodology is not responsive to any other. That distinction is straightforward:

> The language is **simple** if no composition in Rosa's set is eligible for more than two genres[15].

Then there are two contrasting results:

> **Theorem 1.** *Rosa can locate compositions in Rosa's set into genres given a simple language in a way that is responsive to each influencer and that satisfies* **class unanimity**, **pair unanimity** *and* **class consistency**.

> **Theorem 2.** *Rosa can locate compositions in Rosa's set into genres only by nominating a single dominant influencer if the language is not simple and is connected and the aggregator that Rosa uses satisfies* **class unanimity**, **pair unanimity** *and* **class consistency**.

The outcomes of the two theorems are very different, but the eligibilities required need be different only in that theorem 1 holds when no composition is eligible for more than two genres, and theorem 2 holds even if only one composition is eligible for three genres. Rosa can 'follow the majority' (section 4.6) whether or not the genres are connected, and so theorem 1 does not specify connection.

Chapter 10 states and proves the theorems in a more formal way that does not draw on the illustrative context of musical genres. Chapter 12 contrasts them with received theorems in the social choice literature.

[15] It would be more accurate to say that the language is simple "as it applies to Rosa's set and the concepts under discussion". That restriction is implicit in all that follows.

4.6 Theorem 1: following the majority

When no composition is eligible for more than two genres, Rosa can use a **majority methodology** to devise her classification. By using the **simple majority methodology**, Rosa locates each composition in the genre favoured by a majority of her influencers with the additional proviso that she must nominate one influencer as tie-breaker if she defers to an even number of influencers. This methodology satisfies all of the deference principles of theorem 1: the only non-trivial argument is that part (i) of ***pair unanimity*** follows because two majority subsets of the influencers cannot be disjoint, so, if Rosa co-locates two compositions, then at least one influencer co-locates them. An aggregator that is based on the majority methodology is responsive to each influencer because there are circumstances in which an influencer is marginal to some allocation - or marginal to invoking the additional weight of the tie-breaker.

Class consistency prevents Rosa from breaking ties using the toss of a coin because a different coin-toss can change Rosa's location of a composition even though no influencer relocates that composition. The tie-breaking method must involve the location stated by some influencer, who has a greater weight in forming Rosa's opinion than does any other influencer. In that circumstance, Rosa bases her aggregator on a **weighted majority methodology**.

If Rosa nominates influencer 1 as tie-breaker when she has four influencers in total, she calculates her majorities using weights[16] of 0.4 for influencer 1 and 0.2 for each of the other three. Then any three influencers form a majority (having total weight of 0.6 or 0.8), and if their views are equally split, the pair that includes influencer 1 prevails (by 0.6 to 0.4).

Rosa must nominate the same influencer as tie-breaker for her consideration of the location of every composition. If she nominates different tie-breakers for dif-

[16] Weights are standardised so that they sum to 1 across all the influencers.

ferent composition (so using different weights for different location decisions), she faces the possibility illustrated in table 4.3.

influencer	*Rhapsody*	*Prelude*
1 (tie-breaker for *Rhapsody*)	JAZZ	CLASSICAL
2 (tie-breaker for *Prelude*)	CLASSICAL	JAZZ
3	CLASSICAL	JAZZ
4	JAZZ	CLASSICAL
Rosa	JAZZ	JAZZ

Table 4.3: Different tie-breakers

No influencer co-locates *Rhapsody* and *Prelude*, but the intervention of the distinct tie-breakers entails that Rosa co-locates the compositions in JAZZ, contrary to ***pair unanimity***.

4.6.1 Weighted majorities

This methodology can be extended beyond tie-breaking to involve weighted majorities in which Rosa can allow that she regards some influencers as more authoritative than others. The constraints on the weights that she can use are that no single influencer has a weight in excess of 0.5 (otherwise that influencer is dominant), and that she avoids ties by not assigning a combined weight of exactly 0.5 to any subset of her influencers. The weights define subsets of influencers whose combined weight exceeds one-half. If all members of such a subset agree on the location of a composition, then so also does Rosa's location of that composition.

The weights that Rosa uses must be the same, or almost the same, for all compositions. Rosa cannot, for example, follow the three deference principles if she assigns significantly different weights to her influencers for deciding her location of *Rhapsody* compared to those that she uses in deciding her location of *Prelude*. Suppose that Rosa believes that influencer 1 is an expert in the works of Gershwin,

and influencer 2 is an expert in the works of Bernstein. She has five influencers in total and, for each composition, she gives a weight of 0.4 to her nominated expert, and 0.15 to each of the other four. Table 4.4 gives an example in which Rosa co-locates the compositions in JAZZ, but no influencer co-locates them, violating ***pair unanimity***. The weights that she uses differ sufficiently between the compositions that she violates one of the deference principles.

	Rhapsody		*Prelude*	
influencer	weight	genre	weight	genre
1	0.4	JAZZ	0.15	CLASSICAL
2	0.15	CLASSICAL	0.4	JAZZ
3	0.15	JAZZ	0.15	CLASSICAL
4	0.15	JAZZ	0.15	CLASSICAL
5	0.15	CLASSICAL	0.15	JAZZ
Rosa		JAZZ		JAZZ

Table 4.4: Different weights

4.7 Theorem 2: a dominant influencer

Theorem 2 raises significant issues. Given the deference principles, Rosa must follow only one dominant influencer even if that influencer disagrees with all other influencers on the location of every composition in Rosa's set.

The theorem does not specify which influencer is dominant: Rosa can nominate any one of them. In the example on page 64 there are in principle c4.5 million different aggregators of which only 7 remain if Rosa uses all the deference principles of theorem 2. These are the methodologies in which Rosa's classification of the compositions is, for every profile, identical to that of a single influencer.

Rosa can avoid this conclusion only by abandoning one or more of the deference principles. ***Class unanimity*** and ***pair unanimity*** guarantee that Rosa has support

from at least one of her influencers in locating or co-locating compositions. Without this support, her reasoned responses to the questions "why do you (not) locate x in G?" and "why do you (not) co-locate x and y?" involve criteria that include "I disagree with the unanimous opinion of the influencers because I have some other evidence, or because they are not really my influencers". Such a response moves beyond deference. Rosa's 'other evidence' is likely to include her own analysis of the characteristics or properties of the compositions so that her opinion of what it takes to be in some genre, or to be co-located, is different from that of all those to whom she defers.

Likewise, if Rosa decides to depart from ***class consistency***, she needs to justify, for example, why a *ceteris paribus* switch by one influencer of *Rhapsody in Blue* from CLASSICAL to JAZZ leads her to switch *The Sidewinder* from JAZZ to BLUES. The likely answer is that the increased support for locating *Rhapsody* in JAZZ leads Rosa to reconsider the properties needed to locate a composition in JAZZ and decide that *Sidewinder* fits less well in JAZZ than she previously thought. Once again she uses her own evaluation of the compositions to supplement the evidence from the influencers, and we exclude the use of such additional evidence in this chapter (but not in chapter 6).

4.8 Summary

Theorems 1 and 2 form the fundamental conclusion of this dissertation. Difficulties arise for the aggregation of classifications when the language is not simple. Later chapters include variations and developments.

The chapter shows that some simply defined methodologies (majority, preponderance and modal) can have potentially undesirable implications. This leads to the question whether any aggregator can avoid violating some straightforward deference principles. The main deference principles introduced here are that Rosa follows unanimous views about the location and co-location of compositions in

genres, and that her view of the location of composition x depends only on her influencers' statements of the location of x

If Rosa follows these deference principles, there is a fine line between circumstances in which she can respond to every influencer, and circumstances in which she cannot compromise at all between her influencers, because one is dominant. This difference arises according to whether or not any composition in Rosa's set is eligible for three or more genres. Rosa has a fully responsive aggregator only when the language is simple. Proofs are given in chapter 10.

So, theorems 1 and 2 entail that Rosa cannot reach an acceptable compromise between her influencers if there is too much scope for disagreement between them. Rosa is not able to summarise the evidence that she has in an acceptable way. Although the influencers use a language to which they all conform, a non-simple language gives sufficient scope for different usages of concepts that Rosa cannot learn from them in a consistent and responsive way. There are limits to the extent to which it is possible to extend the idea of deference in the formation of concepts from that which was introduced by Burge, Fodor and others. It is unrealistic to suppose that people agree about every genre classification, even within conformity to the language, but there are limits to the extent of disagreement that Rosa can easily accommodate.

Chapter 5

Variations on the theme of deference

5.1 Overlapping concepts

In the formulation of chapter 4, the influencers and Rosa are all constrained to specify that each composition is in exactly one genre. But concepts might overlap, and so an influencer who is asked whether *Rhapsody in Blue* exemplifies JAZZ or CLASSICAL might say "both". This is a possible response when the influencer can see features of *Rhapsody* that fit well with compositions, perhaps archetypes, in JAZZ, and other features that fit well with compositions in CLASSICAL.

If this constraint is relaxed, eligibilities do not change. For example, the two concepts for which *Rhapsody* and *Prelude, Fugue and Riffs* are eligible remain JAZZ and CLASSICAL, but there are now three classes *jazz-only*, *classical-only* and *classical-and-jazz* in which influencers and Rosa might locate *Rhapsody* or *Prelude*. In this formulation, *classical-and-jazz* is not a hybrid or fusion genre in its own right: if it were, it would be a third genre and the two compositions would be eligible for all three, so that the language is not simple. Also, the classes are not identical to the

extensions of concepts: JAZZ is made up of *jazz-only* and *classical-and-jazz*; CLASSICAL is made up of *classical-only* and *classical-and-jazz*.

As long as the language is simple, Rosa can continue to use a majority methodology. For example:

- Rosa locates *Rhapsody* in JAZZ just if a majority of her influencers[17] locate *Rhapsody* in JAZZ - because they locate *Rhapsody* in *jazz-only* or in *classical-and-jazz*;
- Rosa locates *Rhapsody* in CLASSICAL just if a majority of her influencers locate *Rhapsody* in CLASSICAL - because they locate *Rhapsody* in *classical-only* or in *classical-and-jazz*.

This allows that Rosa can locate *Rhapsody* in both genres provided that there is a majority both for its location in JAZZ and CLASSICAL. If both majorities occur, then at least one influencer has located *Rhapsody* in *classical-and-jazz*. So Rosa's deference to unanimous views extends to ensure that if no influencer locates a composition in both of the two genres for which it is eligible, then Rosa does not locate it in both genres.

It is possible that Rosa locates *Rhapsody* in both genres even though only one influencer locates it in both as shown by the simple example of table 5.1. Rosa follows the majority view on each genre location, and so she locates *Rhapsody* in both genres. There is an even split of opinion amongst her influencers, and the possibility of locating *Rhapsody* in both genres avoids the need for Rosa to choose a single genre, which would be somewhat arbitrary given the even split of opinion amongst her influencers.

[17] Including a tie-breaker if that is relevant.

influencer	location of *Rhapsody*
1	JAZZ & CLASSICAL
2,3,4,5	JAZZ
6,7,8,9	CLASSICAL

Table 5.1: Overlapping genres

The presence of three classes (but only two genres) does not introduce the difficulties of theorem 2 that arise when a composition is eligible for three genres. With two genres and three classes, Rosa's aggregator[18] is that:

- Rosa locates *Rhapsody* in *jazz-only* if a minority of her influencers locate it in *classical-only* or in *classical-and-jazz*;
- Rosa locate *Rhapsody* in *classical-only* if a minority of her influencers locate it in *jazz-only* or in *classical-and-jazz*;
- otherwise, Rosa locates *Rhapsody* in *classical-and-jazz*.

Classical-and-jazz is the residual class in which Rosa locates *Rhapsody* if neither of the other rules apply.

This contrasts with an example in which *The Sidewinder* is eligible for JAZZ, BLUES and SOUL and Rosa does not regard one genre as a residual location. In that case, theorem 2 can apply: any aggregator that does not involve a dominant influencer violates at least one of the deference principles. For example, table 4.1 shows that an aggregator that is based on the preponderance methodology can violate ***pair unanimity***.

[18] Redefined to allow that a composition can be located in more than one genre.

5.2 Extending concepts

There are many compositions with which Rosa is unfamiliar. 'Rosa's set' - with which she is concerned - is a subset of the universe of all compositions and, at any time, the universe is finite, albeit very large. An important part of the development of thinking about concepts arises when Rosa discovers a composition (called *New to Rosa*) that she has not previously considered. The language defines the eligibilities of *New to Rosa* and we suppose that Rosa's influencers already have opinions about the location of *New to Rosa* within the genres for which it is eligible. When Rosa expands her set, the influencers continue to give their opinions about the genre locations of compositions that were already in Rosa's set. So the expansion of Rosa's set does not change any influencer's view about the location of compositions previously known to Rosa. This can contrasted with the possibility that an influencer who discovers a new composition might want to change the location of some other compositions because the new discovery impacts his view about, for example, what jazz music is.

Theorems 1 and 2 can be applied (according to the eligibilities) to Rosa's set both before and after its expansion to include *New to Rosa*. There are three scenarios, given the deference principles:

1. Before and after the expansion, no composition is eligible for more than two genres

2. Before the expansion at least one composition in Rosa's set is eligible for more than two genres.

3. Before the expansion, no composition is eligible for more than two genres, but *New to Rosa* is eligible for more than two genres.

The assumption that there is no change in any influencer's location of compositions already in Rosa's set, and ***class consistency*** can be combined to entail that Rosa's location of every composition that was previously in Rosa's set does not

change when she includes *New to Rosa*. The first two scenarios are straightforward. In scenario 1, Rosa can continue to devise her concepts using a majority methodology and apply it before and after the expansion. In scenario 2, Rosa has chosen her dominant influencer before the expansion, and continues to defer only to that influencer to determine her location of *New to Rosa*.

Scenario 3 is more problematic, because Rosa can use a majority methodology before she includes *New to Rosa*, but cannot do so after if she adheres to the deference principles of theorem 2. After the expansion of Rosa's set, she must nominate a dominant influencer, but there will be circumstances in which her dominant influencer was previously not in the majority group for the location of (say) *Rhapsody in Blue*, leading to a change in Rosa's location of *Rhapsody*. To avoid this incompatibility with ***class consistency*** as we have applied it to the expansion of Rosa's set, Rosa must nominate a dominant influencer even before she considers *New to Rosa* - leading to the uncomfortable conclusion that Rosa defers to a dominant influencer now if it is possible that, at some time in the future, she will include a composition that is eligible for more than two genres. She would need to take this step even though she has no idea of the identity of *New to Rosa*, or even whether she will at some stage become familiar with such a composition.

Technically, this precautionary nomination of a dominant influencer can be avoided by allowing that Rosa's location of *Rhapsody* can change if there is a change in the set of compositions under consideration, even though no influencer's location of *Rhapsody* is impacted by the inclusion of *New to Rosa*. Then ***class consistency*** does not apply when Rosa's set expands. Rosa might form her concepts in this way, but is most likely to do so if the introduction of the additional composition changes her view of the properties associated with JAZZ. If Rosa previously located *Rhapsody* in CLASSICAL but then locates *New to Rosa* in JAZZ, she might then come to think that *New* and *Rhapsody* have sufficient similarities that they should be co-located. This sequence moves Rosa away from reliance only on deference to include a consideration of properties (chapter 6).

5.3 Defining a dialect

The term 'dialect' refers here to the way in which a sub-community uses the language (section 1.4). If the definition of the dialect of the sub-community depends on its members' usages, similar considerations can arise to those that are relevant to Rosa's development of concepts through deference to her influencers. The deference principles transform into conditions or restrictions on the way in which the dialect is defined, given the individuals' usages. An analogue to ***class unanimity*** seems to be uncontroversial: if no member of the sub-community locates *The Sidewinder* in BLUES then the defined dialect also excludes *Sidewinder* from BLUES; if every member locates *Sidewinder* in SOUL then so does the defined dialect. Likewise, ***pair unanimity*** gives sufficient conditions for the dialect to identify similarity by co-locating two compositions - and similarly for the identification of dissimilarity. ***Class consistency*** allows that the genre location of a composition in the dialect relies only on the locations of that composition in the usages by the group members.

Then theorem 1 indicates that the dialect can be defined on the basis of majority opinions, provided that disagreement within the group is limited to a simple language. Theorem 2, on the other hand, establishes that the dialect to be used must be determined by a single dominant member of the sub-community when the language is not simple. The members of the sub-community might regard this solution to their need for a commonly agreed use of terms as leadership or as dictatorship - and their attitude might depend on the nature and purpose of their sub-community.

5.4 Summary

This chapter extends the analysis of deference in chapter 4 by allowing that concepts can overlap. The third possibility for an influencer's location of, say, *Rhapsody*

in Blue (that it is in both JAZZ and CLASSICAL) does not define a third genre, and so does not introduce the potential difficulties that lead to theorem 2.

The recognition that Rosa's set is a proper subset of the universe of compositions raises some problems of consistency if Rosa's set expands over time. There are circumstances in which Rosa must either nominate a dominant influencer even before Rosa's set contains any compositions that are eligible for more than two genres, or accept that she must change her location of some compositions when that expansion occurs.

The third part of this chapter applies the analysis of deference to the alternative context of defining a dialect for a sub-community (of a larger language community) whose usages of the language differ, but who - for some reason - define a coherent group that has, or needs to have, common usage. The conclusions of theorems 1 and 2 remain. If they disagree too extensively, the members of the sub-community must accept that their dialect is the same as the usage of one member.

Chapter 6

Evaluating properties

6.1 Introduction

This chapter explores an alternative approach to the formation of concepts. Instead of deferring to others, Rosa makes her own evaluation of whether compositions are sufficiently similar that she co-locates them. This process is likely to be applied when she is familiar enough with compositions and genres that she has identified a composition that she definitely locates in each genre, possibly by recognising archetypes through her earlier deference to others. So, for any composition x, her co-location decisions involving x and the various archetypes entail her location decision for x.

This alternative approach to locating compositions in genres also involves reconciling conflicting evidence. A composition is eligible for more than one genre because it has some properties that support each of its eligible locations, and Rosa needs to evaluate their relative influence. The main question - clearly suggested by theorem 2 - is whether Rosa faces potential difficulties when the evidence concerning properties is sufficiently diverse.

The exploration of this question requires the formalisation of the relation between Rosa's perception of the properties of compositions and the outcome that she reaches on their genre locations. This formalisation is done in two ways by using **property classifications** (section 6.2) and using **property rankings** (section 6.3). Both of these approaches initially suggest analogous difficulties to those of theorem 2. If the evidence is sufficiently diverse, and defensible conditions are imposed on the way in which she aggregates the evidence, Rosa is led to a 'dictatorial' outcome in which her location decision for every composition depends only on a single property. This is plainly unsatisfactory in many contexts, but section 6.4 introduces a significant additional consideration that moderates this possible outcome by questioning the applicability of unanimity and consistency conditions. It is much more difficult to moderate these conditions in the deference approach (section 6.5).

Without explicit deference to others, Rosa has no mechanism through which she can learn to conform to the language. Her consideration of properties might lead her to fail to conform to eligibilities. As an (extreme) example, suppose that Rosa equates JAZZ with music that is syncopated, and takes little notice of other properties. Then she might be tempted to locate CPE Bach's *Magnificat* [76] in JAZZ, even though it is unlikely that any musicologist or critic would agree with that location. Section 6.6 reflects the possibility that Rosa might defer at least to the extent of learning conformity to the language, but that she also considers properties when locating compositions that are eligible for more than one genre. As the section shows, even this hybrid approach can lead to difficulties when the evidence that Rosa uses is sufficiently diverse.

6.2 Property classifications

Rosa has a list of properties that she uses to evaluate compositions. To do this, she devises **property classifications** that give her view of the properties that support the location of each composition in each genre. Her challenge is to find a way of

devising an outcome classification of the compositions into genres on the basis of these potentially diverse property classifications.

When Rosa defers to others, she devises her classification of compositions into genres by aggregating a profile of classifications, one for each of her influencers. In this section, Rosa's challenge involves aggregating a profile of property classifications to determine her view of the genre location of each composition. The main question is whether the deference principles of section 4.3 can be translated into plausible conditions on the form of the aggregator of property classifications - and, if so, how theorems 1 and 2 apply when Rosa uses property classifications.

It seems clear how Rosa will respond when, in her assessment, no property supports the inclusion of a composition in genre G. Using a principle analogous to ***class unanimity*** in such circumstances, Rosa does not locate that composition in G. By extension, she locates a composition in G if every property supports that location.

Rosa's consideration of properties is likely to lead her to co-locate $x\&y$ if she thinks that they are 'sufficiently similar', subject to her recognition that statements of sufficient similarity must be transitive. The property classifications that determine her location decisions give maximum support to co-location if $x\&y$ are sufficiently similar according to every property, so that they are co-located in every property classification. Likewise, there may be no property classification in which $x\&y$ are co-located leading Rosa not to co-locate $x\&y$. So Rosa is likely to use a condition that is analogous to ***pair unanimity*** when she aggregates property classifications.

If Rosa changes her mind *only* about the properties that support her in locating composition x, the analogous restriction to ***class consistency*** entails that she might re-locate x but does not relocate any other composition. Her genre location of a composition is independent of the properties that support the location of any other composition.

The main difference between the specification of theorems 1 and 2 is that, in the former, the language is simple. When Rosa forms her concepts by aggregating property classifications, it might be that no composition in Rosa's set is supported by its properties for inclusion in more than two genres. This would occur if, for example, the properties of *Rhapsody in Blue* all support its location either in JAZZ or in CLASSICAL. Then Rosa would not locate *Rhapsody* in any other genre.

Theorem 1 then shows that Rosa could use a majority methodology and locate x in G just if a majority of properties support that location. If necessary, she could give greater weight to one property in order to break ties. Whilst this aggregator is consistent with the analogues to the deference principles, it might be felt to be very mechanical as a way of representing Rosa's thought processes when she is considering properties. She might, for example, think that tonality is a very important property for some compositions, whilst for others she should give greater importance to rhythmic structure.

	Rhapsody		*Prelude*	
property	weight	genre	weight	genre
π_1	0.4	JAZZ	0.15	CLASSICAL
π_2	0.15	CLASSICAL	0.4	JAZZ
π_3	0.15	JAZZ	0.15	CLASSICAL
π_4	0.15	JAZZ	0.15	CLASSICAL
π_5	0.15	CLASSICAL	0.15	JAZZ
Rosa		JAZZ		JAZZ

Table 6.1: Differential weights on properties

As the example of table 6.1 (table 4.4 applied to properties) shows, using different weights for devising her locations for two compositions can lead to failure to comply with a principle analogous to ***pair unanimity***. Rosa co-locates *Rhapsody* and *Prelude*, but this co-location is supported by no property.

If properties of some composition support its location in more than two genres - so that in effect the language is not simple - it is highly likely[19] that Rosa can satisfy the analogues to the deference principles only by nominating a single 'dominant property' which she uses alone to establish her genre classification. In those circumstances, she ignores all properties except one - a conclusion that seems to be even less acceptable than the nomination of some dominant influencer. It seems to be particularly perverse for Rosa to consider initially that several properties are relevant and then to conclude that principles of unanimity and consistency lead her to change her mind. But that is a potential implication of theorem 2 in this context.

6.3 Property orderings

An alternative formulation of Rosa's thought processes involves a set of **property orderings** for each composition. Given a composition x, she might argue that its rhythmic structure most strongly favours the location of x in CLASSICAL, gives some support to the location of x in JAZZ, marginally supports the inclusion of x in BLUES and gives no support to its inclusion in any other genre. More generally, for each composition and each property, Rosa establishes an ordering (possibly allowing 'ties') of the genres from the most to the least strongly supported. She then aggregates the property orderings for x to establish a 'winner' which is her genre location of x.

This process has structural similarities that are similar to contexts to which Arrow's theorem applies. Arrow's theorem is most frequently applied to contexts such as voting in which the orderings are supplied by individuals, but it can formally be applied to the aggregation of other orderings. Rubinstein & Fishburn (1986) cite an example that

[19] We have not specified a parallel to linked genres, but a diverse set of compositions and properties is likely to suffice to give an analogue to theorem 2.

> assumes that a person is to rank songs according to her tastes. …such a preference may be based on …factors such as originality, rhythm, emotional content, and chord structures, so that her holistic ranking can be viewed as an aggregation…

Also, there has been a discussion following Okasha (2011), including Morreau (2015), Stegenga (2015) and Okasha (2015) who discuss how Arrow's theorem might apply to orderings of scientific theories according to the 'properties' considered by Kuhn (1969, 1977). These relate to the accuracy, consistency, scope, simplicity and fruitfulness of alternative theories and under Arrow-like conditions, a single property might dominate all others.

The three main conditions used in most proofs of Arrow's theorem are that the aggregator satisfies conditions that involve unanimity and independence (of irrelevant alternatives) and that the domain of the aggregator is unrestricted over some triple. If these hold, the outcome is dictatorial.

For Rosa, unanimity in this context entails that if, for example, JAZZ is above[20] BLUES in the every property ordering for *The Sidewinder*, then Rosa does not locate *Sidewinder* in BLUES. Independence entails that if, for example, Rosa is considering *Rhapsody in Blue* and *Prelude, Fugue and Riffs*, and each property ordering for *Prelude* places CLASSICAL above JAZZ just if the same property ordering for *Rhapsody* places CLASSICAL above JAZZ, then Rosa does not locate one of the two compositions in CLASSICAL and the other in JAZZ. In general, Rosa decides whether composition x is more appropriately located[21] in CLASSICAL than in JAZZ depending only on the order of CLASSICAL and JAZZ in her property orderings for x.

[20] Equal rankings can be accommodated as in the contrast between weak Pareto conditions and strong Pareto conditions in welfare economic applications of Arrow's analysis.

[21] But not necessarily located if she decides that some other genre is even more appropriate as a location for x.

ranking	first	middle	last
1	JAZZ	BLUES	SOUL
2	JAZZ	SOUL	BLUES
3	BLUES	JAZZ	SOUL
4	BLUES	SOUL	JAZZ
5	SOUL	JAZZ	BLUES
6	SOUL	BLUES	JAZZ

Table 6.2: Property orderings with three genres

In the property orderings context, the domain consists of the set of profiles of property orderings for the compositions in Rosa's set. A triple of genres can appear for a composition in any one of the 6 strict orders [22] of table 6.2.

If Rosa's set contains compositions for which her property orderings include all of the possibilities in table 6.2, and she devises her genre locations using the unanimity and independence conditions described above, then Arrow's theorem entails that she must base her locations on a single dominant property. Domain restrictions in this context can avoid this difficult conclusion, but this possibility is likely to rest on the compositions included in Rosa's set. For example, if there is no composition in Rosa's set for which her property orderings are those of rows 4&6 in table 6.2 (in which JAZZ is the lowest ranked genre) her rankings are single-peaked (section 12.2.1). Then Rosa can consistently use a majority methodology to locate all the compositions. But if Rosa's set contains many and varied compositions, such a domain restriction seems implausible, and she needs a different way of avoiding a dominant property[23]. An alternative possibility is discussed in section 6.4.

[22] There are more if equal rankings are allowed.

[23] The specification of domain restrictions that are consistent with the use of a majority methodology is even more complex when there are more than three genres - see for example the example of figure 12.2, page 185.

6.4 Intensity

Arrow's theorem, our own theorem 2, and others in the literature (section 1.3) all indicate that sufficiently extensive diversity in the evidence being aggregated leads to the inevitability of dictatorship when the aggregator satisfies unanimity and independence/consistency conditions. Sections 6.2 and 6.3 indicate that this conclusion can apply when Rosa is considering the properties of compositions as her route to establishing genre allocations using either property classifications or property orderings. This is plainly an unsatisfactory implication because there are many contexts, including musical genres, in which Rosa would not expect that her allocations depend only on a single property out of several[24]. This section recognises that there is a feature of the consideration of properties that does not have a direct analogue in the deference model - namely the consideration of the **intensity** with which Rosa feels that properties support the genre location of a composition. This consideration leads to plausible reservations about the use of unanimity and consistency conditions.

For example, if she uses property classifications, there seems to be no barrier to Rosa thinking that, in her view, property π_1 *strongly* supports the location of *Rhapsody in Blue* in JAZZ and that property π_2 *marginally* supports its location in CLASSICAL. This consideration of intensities might lead Rosa to violate the analogous conditions to ***pair unanimity*** or ***class consistency***.

	compositions		
property	x	y	z
π_1	JAZZ++	JAZZ	CLASSICAL
π_2	CLASSICAL	CLASSICAL++	JAZZ++
++ denotes strong support its absence denotes marginal support			

Table 6.3: Intensity example 1

[24] Single property contexts are explored in chapter 7 and a measure of intensity in section 7.2.1.

In the example of table 6.3, without considering intensities, ***pair unanimity*** entails that Rosa co-locates $x\&y$, and does not co-locate $x\&z$ or $y\&z$. But when she allows for the intensity of support, Rosa can think as follows:

- Property π_1 strongly supports JAZZ as the location for x, marginally supports JAZZ as the location for y and marginally supports CLASSICAL as the location of z.
- Property π_2 marginally supports CLASSICAL as the location of x, strongly supports CLASSICAL as the location of y and strongly supports JAZZ as the location of z.

Then it is reasonable to suppose that Rosa locates $x\&z$ in JAZZ and y in CLASSICAL contrary to ***pair unanimity***.

	location of x	
property	before	after
π_1	JAZZ++	JAZZ
π_2	CLASSICAL	CLASSICAL++

Table 6.4: Intensity example 2

Table 6.4 gives a simple example in which Rosa assigns the same property locations to x but changes the intensities. The change in intensities could support her in locating x in JAZZ before the change and in CLASSICAL after the change, contrary to ***class consistency***.

If she allows for the intensity with which she feels that a particular property supports a genre location, Rosa might behave in ways that violate conditions that are analogous to those that lead to a dominant individual in theorem 2 - and so she is not inevitably led to nominate a dominant property.

A similar conclusion can be reached if Rosa's use of properties is based on property orderings. For example, the analogue to Arrow's ***independence of irrelevant***

alternatives might not hold given the orderings and intensities of table 6.5, where $G, G'\&G''$ are three genres.

	composition x	composition y
property	ordering	ordering
π_1	$G \blacktriangleright\blacktriangleright\blacktriangleright G' \blacktriangleright G''$	$G \blacktriangleright G' \blacktriangleright\blacktriangleright\blacktriangleright G''$
π_2	$G' \blacktriangleright G'' \blacktriangleright G$	$G' \blacktriangleright\blacktriangleright\blacktriangleright G'' \blacktriangleright G$
π_3	$G'' \blacktriangleright G \blacktriangleright\blacktriangleright G'$	$G'' \blacktriangleright G \blacktriangleright G'$
Rosa	x in G	y in G'
$\blacktriangleright \equiv$ "ranked above" intensity increases with additional $\blacktriangleright$		

Table 6.5: Intensity example 3

Each property ordering is the same for x and for y, but the different intensities could justify the different outcomes that Rosa locates x in G and y in G' in violation of the analogue to ***independence of irrelevant alternatives***.

6.5 Intensities in deference

The obvious follow-up question to section 6.4 is to examine whether Rosa might use an intensity argument when she defers to others, so that she can avoid the conclusion of theorem 2.

The response lies in the nature of the evidence that Rosa uses in the two approaches. When she considers the properties of performances, the question of how strongly she feels is internal to Rosa. She uses her own understanding of the intensity of feeling and of comparisons of the intensity with which she feels that each property supports a particular genre location or co-location. If she attempts to use an intensity argument when she defers to others, she needs to find out how intensely her influencers feel about the locations that they state. As well as the evidence-gathering challenge of getting evidence from her influencers about the

intensities of their views, she faces the need to compare the intensity statements of different influencers. For example, if Rosa asks them to indicate the strength of their support for locating *The Sidewinder* in BLUES on a scale that runs from 1 to 5, she has no standard which she can use to compare the scores stated by each influencer.

A similar challenge of interpersonal comparisons arises in attempts to aggregate cardinal utilities in ethics or in welfare economics[25]. Does 'one util' entail the same intensity for two individuals? There is a similar difficulty with opinion surveys in which people are asked to locate their views on a topic on a 5-point scale from 'strongly disapprove' to 'strongly approve', or to add scores that try to represent the aptitudes of job candidates. How can those scores be aggregated in a way that recognises that individuals interpret the scale differently?

Rosa's challenge to assess the intensity with which influencers hold their views does not involve the same considerations as her ability to assign different weights to influencers in applying a weighted majority methodology. The example of table 6.1 shows that this interpretation of differential influence can lead to a violation of ***pair unanimity***.

An alternative interpretation of intensity when Rosa defers to others arises in section 7.2 in a graded context. Then Rosa can meaningfully establish 'how far apart' are two objects in the opinion of each influencer - but this seems not to be realistic in the consideration of genre allocations.

6.6 Mixing deference and properties

Deference to influencers who conform to the eligibilities defined by the language, and in particular the use of ***class unanimity***, enables the possibility that Rosa is corrected if she uses a concept in a way that does not conform to the language.

[25] A major stimulus to the development of social choice theory was the recognition that using ordinal scales of utility was less problematic than the use of cardinal utility.

There is no such clear route to correction of non-conformity if Rosa uses only her own evaluation of properties. As noted in section 6.1, is possible that some property gives support to a genre for with the composition is ineligible, and so that location is not immediately excluded by Rosa.

This section considers that Rosa might use two sources of evidence:

- a classification that is based on Rosa's evaluation of the properties of the compositions, plausibly including the recognition of different intensities.
- the classifications stated by each influencer.

The former is Rosa's ***ex ante* classification** that she holds before she considers the possibility of deferring to the evidence from others. She might then modify this through deference to her influencers to form her ***ex post* classification**.

Rosa's response is likely to follow a methodology in which she does not modify her *ex ante* location of a composition if sufficient experts agree with her about that location. She modifies an *ex ante* location if sufficient experts disagree with her and agree with each other about some alternative location.

In following this sort of methodology, Rosa is likely to want to be consistent with some principles[26]. A principle analogous to ***class unanimity*** remains reasonable: if Rosa's *ex ante* classification does not locate x in G, and no influencer locates x in G then Rosa's *ex post* classification does not locate x in G. By extension, if Rosa's *ex ante* classification and every influencer locates y in G then Rosa's *ex post* classification locates y in G.

If Rosa is seeking support for the *ex post* co-location of two compositions, she would expect at least that her own *ex ante* classification or that of some influencer would co-locate them. Likewise, if Rosa is seeking support for not co-locating two compositions *ex post*, she would expect at least that her own *ex ante* classification

[26] These remain deference principles if we allow that Rosa defers to herself as she is represented by her *ex ante* classification.

or that of some influencer would not co-locate them. So Rosa follows a principle analogous to ***pair unanimity***.

Finding a principle that is analogous to ***class consistency*** is arguably more problematic. Suppose that, before Rosa re-evaluates the evidence, she decides that her *ex ante* classification strongly supports the location of *Rhapsody in Blue* in JAZZ on the basis of its properties, and that this strong support just outweighs the evidence that a large proportion of her influencers locate *Rhapsody in Blue* in CLASSICAL. If Rosa re-evaluates the properties and decides that her *ex ante* classification marginally supports the location of *Rhapsody in Blue* in JAZZ, this lesser support might no longer, in Rosa's thinking, outweigh the numerical weight of influencers' support for locating *Rhapsody in Blue* in CLASSICAL. Although she does not know the intensity with which any others hold an opinion, she has a proxy for the overall intensity of their opinion in the number of influencers who hold each view. If this relative intensity argument is ignored, a principle analogous to ***class consistency*** can apply in this mixed methodology.

Rosa has enriched her evidence base for determining her *ex post* classification, but might continue to face similar challenges that arose when she relied only on deference.

6.6.1 Simple language

If no composition is eligible for more than two genres, Rosa can use a fully responsive methodology to determine her *ex post* classification. Rosa might decide that her *ex post* location of a composition differs from her *ex ante* location if at least some fraction of her influencers disagree with her[27]. For example, if Rosa has 9 influencers, and uses a methodology through which she changes her location of a composition if more than 6 influencers disagree with her, she is effectively using a weighted majority methodology in which her own *ex ante* classification has a

[27] Rosa could apply weights to the influencers complicating the exposition but not adding significantly to the conclusion.

weight of 0.28, and each influencer has a weight of 0.08. Rosa and 3 influencers have a combined weight of 0.52; 7 influencers have a combined weight of 0.56. Each such grouping can determine the outcome if they agree. So, with a simple language, Rosa can establish a trade-off between the opinions of others against her own *ex ante* views in a straightforward way. In particular, Rosa will be corrected if, *ex ante*, she locates a composition in a genre for which it is not eligible because all her influencers contradict her *ex ante* location.

6.6.2 Non-simple language

Theorem 2 indicates that no such trade-off is possible when one or more composition is eligible for three or more genres. Then if Rosa uses an aggregator that satisfies ***class unanimity***, ***pair unanimity*** and ***class consistency***, she must nominate a single influence as dominant. She cannot use a methodology that involves both her own classification and the opinion of any influencer. Either she ignores the opinions of others, or she ignores her own *ex ante* evaluation of the properties of compositions by nominating a single dominant influencer. The conclusion that significant diversity in her evidence base causes difficulties for Rosa is not dispelled by the use of both her own evaluation and the opinions of experts.

6.6.3 Mutual interaction

The analysis of equilibrium in games or in economic behaviour suggests the possibility that mutual interaction within a community might lead its members to, or towards, agreement on the way in which concepts are used. If community members are influenced by one another, it seems plausible that outliers are persuaded towards classifications that are supported by a significant number of others. Such an outlier has an idiosyncratic *ex ante* classification that he adjusts when he assesses the classifications of others. The dialect of the community might be defined using a majority aggregator because any equilibrium set of classifications shows insufficient diversity to lead to a dictatorial outcome. If the language is simple, then

the locations defined by each individual in the community are likely to converge towards the majority view, each starting from her own *ex ante* classification.

However, if the language is not simple, and the relevant deference principles hold, such convergence cannot happen. Indeed, having devised her own *ex ante* classification, it is arguable that no individual wants to ignore it completely. Then each individual must identify her *ex post* classification with her own *ex ante* classification. There is no scope for mutual interaction[28]. Every member of the community defers only to herself, and so continues to use her *ex ante* classification. This removes the possibility that mutual interaction might reduce the extent of diversity of views. Each individual persists in her initial views.

6.7 Summary

This chapter examines the possibility that Rosa devises her genre locations of compositions by her own internal consideration of the properties of compositions. *Rhapsody in Blue* is eligible for JAZZ and for CLASSICAL because there is 'disagreement' about the implications of different properties.

The evidence from properties can be modelled in at least two ways, and given a re-interpretation of the deference principles of chapter 4 there are circumstances in which an equivalent of theorem 2 holds. Rosa must base her genre location of every composition in Rosa's set on the evidence from only a single property. There is too much disagreement between the evidence from properties to allow her to find a compromise.

Rosa's consideration of properties is internal to her, and so it is plausible to suggest that she can avoid the difficult conclusion of theorem 2 by taking account of her feelings about the intensity with which properties support genre locations. Then the conditions that are analogous to deference principles can reasonably be vio-

[28] In a different context, List (2011) recognises this consequence of 'dictatorship' results.

lated, and Rosa can avoid nominating a dominant property. It is much more difficult to find and assess evidence of intensity when Rosa is deferring to her influencers, because there is no clear common scale from which can Rosa derive intensities and aggregate them.

Rosa might combine consideration of properties with deference to her influencers, and there are circumstances in which she can use a variation of the majority methodology to adjust her own property-based view if sufficient influencers disagree. But if there is too much disagreement, so that theorem 2 holds, Rosa must either rely only on, or completely abandon, her property-based view in the face of disagreement from her influencers.

Chapter 7

Graded contexts

7.1 Introduction

This chapter reverts to the consideration of deference alone. In the analysis of chapter 4, influencers can locate two compositions, such as *Rhapsody in Blue* and *Prelude, Fugue and Riffs*, in any of four ways: both in JAZZ, both in CLASSICAL, or one in each genre. The approach of chapter 6 suggests that this four-way possibility arises because each composition has several properties, some of which suggest that JAZZ is the most appropriate location, and others suggest CLASSICAL. An individual can combine these properties to reach any of four classifications of *Rhapsody* and *Prelude*, depending on the individual's views of their relative importance.

This chapter concerns objects that differ in only one property, and that can be **graded** because they can be placed in order according to that unique relevant property. This in itself entails a restriction on the possible classifications of the objects under consideration, and on the extent of disagreement. For example, if individuals agree about the grading of trees by height, it is not possible for one to locate tree x in TALL and y in SHORT, whilst another locates x in SHORT and y in TALL - assuming that they both understand that the use of 'tall' and 'short' in their language entails that anything tall is of greater height than anything short.

In the tall trees example, there is a straightforward underlying measure of height that might be used to create an average so that TALL consists only of trees that are above average height. In a specified forest of trees, this leaves no scope for disagreement after measurements have been taken. But there are some ordinal gradings that are not easily susceptible to measurement, and where averaging is meaningless. Arguably, successive days with noon temperatures of 30°C and 10°C are not the same as successive days both at 20°C although the average temperature is the same. Non-linear transformations of temperature scales are possible and preserve the grading information but render averages difficult to interpret.

Our illustrative context is graded and the scale is not open to meaningful linear transformations. It involves the **spiciness** of a set of **dishes** (in the sense of food prepared as part of a meal). The Scoville scale assigns numbers to capsicums according to their spiciness or "heat", and implicitly to food prepared with them, but the use of this measure requires sophisticated instrumentation for High Pressure Liquid Chromatography [11]. We can suppose that people can agree about the ranking of dishes by spiciness without having access to measurement - or if they do have access, averaging is unhelpful[29]. The challenge for Rosa is to use an ordinal grading scale and to devise concepts such as MILD, MODERATE and HOT given the views of the influencers to whom she defers.

The analysis here concerns dishes that are graded according to some single intrinsic property of the dishes that is not separately evaluative. People might disagree about whether they prefer hotter curries to milder curries, or about whether the relation 'more delicious than', or 'preferable to', is positively or negatively correlated with spiciness. Plunkett & Sundell (2013) develop an analysis of evaluation that explores the use of a relation such as 'better than' or a term such as 'good' when people

[29]The hottest chilli peppers are around one million on the scale; bell peppers are close to zero, but there is nothing helpful beyond arithmetic truth in the statement that a 500k chilli (very spicy) is an average of the hottest and mildest. There is no guarantee that two 500k peppers can be used in place of one at one million and one at zero without a major impact on the outcome of a recipe.

have radically opposed views of how it applies[30]. Disagreement about these sorts of evaluations is beyond our scope.

7.2 Grading, ranking and boundaries

The influencers agree that dishes in Rosa's set are graded by a transitive binary relation $\succ$: "$y \succ x$" ≡ "y is spicier than x". Each influencer locates each dish into a concept for which it is eligible according to the language, recognising that, for example, Phaal Curry is not eligible for MILD and Chicken Korma is not eligible for HOT. The r concepts (MILD, MILDISH, MODERATE ...) are ranked[31] in the obvious way, and each influencer partitions the list of dishes into r non-empty convex subsets, one for each concept, so that each influencer makes use of every concept.

The term **neighbouring** is used in its obvious sense within Rosa's set. $y \succ x$ are neighbouring dishes just if there is no z in Rosa's set such that $y \succ z \succ x$.

An alternative and equivalent approach to Rosa's challenge in a graded context requires that she places[32] boundaries between some neighbouring pairs of dishes. These boundaries show how she uses the concepts. She does this by deferring to her influencers who each state where they place the boundaries. The eligible places for boundaries can be derived from the eligibilities of the dishes defined by the language. For example, if x is eligible to be in MILD and neighbouring y is eligible to be in MODERATE then the boundary between MILD & MODERATE is eligible to be placed between $x\&y$.

[30] A classic, but now contextually difficult, example was used by Hare (1952) who recognised that there is some commonality in the use of 'good' by headhunters, for whom virtue correlated positively with the number of enemy heads that they collected, and missionaries, who took the opposite view. The two groups have opposing definitions of good actions, but both identify 'good' with 'worthy of praise or reward'.

[31] 'Graded' is used for the dishes and 'ranked' for the concepts.

[32] 'Locate' is used for the dishes and 'place' for the boundaries.

The deference principles (section 7.5) used by Rosa are stated in terms of her response to the boundary placements stated by her influencers. This leads to some technical differences from those of the non-graded context that are further explored in chapter 11.

7.2.1 Intensity

Chapter 6 explored the possibility that Rosa can evaluate the intensity with which genre allocations are supported. The main problem in that context is that there is no clear measure of intensity that Rosa can use to aggregate the opinions of influencers to whom she defers. But, when dishes are graded and concepts are ranked, it is possible to attribute an ordinal measure of intensity to the opinions of influencers.

Suppose, for example, that $y \succ z \succ x$ and an influencer locates y in HOT, z in MODERATE and x in MILD. Then he considers that the three dishes are sufficiently different that they are all in different concepts, and that this view of difference is stronger when comparing $y\&x$ than in comparing $y\&z$ or $z\&x$. The influencer places two boundaries between $y\&x$, and one between the other pairs, and the proxy measure of the intensity of the difference between two dishes is the number of boundaries placed between them. With this interpretation:

"i does not change the intensity of the difference between two dishes"

is equivalent to

"i retains the same number of (concept) boundaries between the dishes".

The intensity measure ascribed by an influencer to the difference between two dishes is zero just if the individual co-locates the dishes in a concept.

7.3 Simple and non-simple languages

In the graded context, a simple language is one in which there is no place for which two boundaries are eligible. This is trivially the case when there are only two concepts and hence a single boundary, as illustrated in figure 7.1.

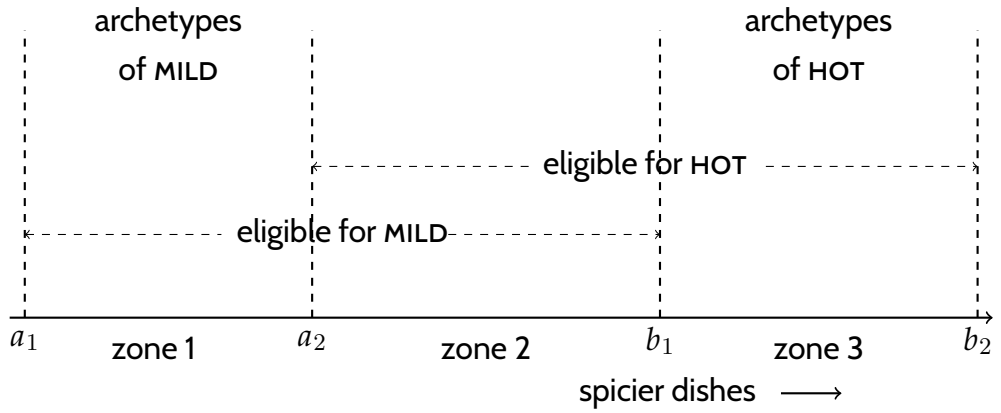

Figure 7.1: Eligibility: two concepts

The language defines the eligibilities, and each influencer places the boundary between MILD and HOT to include all, some or none of zone 2 in MILD with the remainder of zone 2 in HOT. That single boundary is eligible to be placed just below[33] a_2, or anywhere between $a_2 \& b_1$, or just above b_1. So anyone who conforms to the eligibilities of the language regards dishes in zone 1 (a_1 but not a_2) as archetypes of MILD and dishes in zone 3 (b_2 but not b_1) as archetypes of HOT. Dishes in zone 2 (including a_2 and b_1) are eligible for both concepts.

When there are three concepts MILD, MODERATE and HOT, there are two possible configurations. These are shown in figures 7.2 and 7.3.

[33] A boundary is 'just below' dish x if there is no dish between x and the boundary. A similar definition applies to 'just above'.

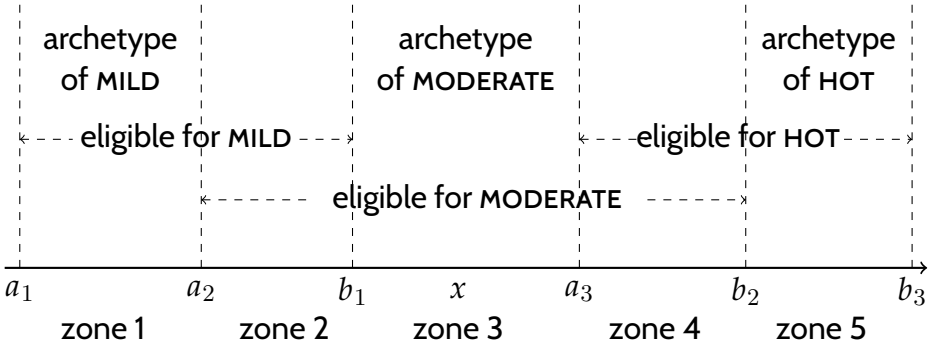

Figure 7.2: Three concepts: simple language

In figure 7.2, the boundary between MILD & MODERATE can be placed just below a_2, between $a_2 \& b_1$ or just above b_1. The boundary between MODERATE & HOT can be placed just below a_3, between $a_3 \& b_2$ or just above b_2. The chapter 4 definition of a simple language is no longer fully adequate. In figure 7.2 no dish is eligible for more than two concepts, but the language is not simple unless 'just above b_1' is distinct from 'just below a_3'. This entails that there must be a dish (x in figure 7.2) in zone 3 that is graded between b_1 and a_3. Then x is an archetype of MODERATE. This is the **archetype assumption** that in a simple language there is an archetype of each concept.

The alternative configuration of figure 7.3 shows that if some dish is eligible for three concepts, there cannot be an archetype of each concept.

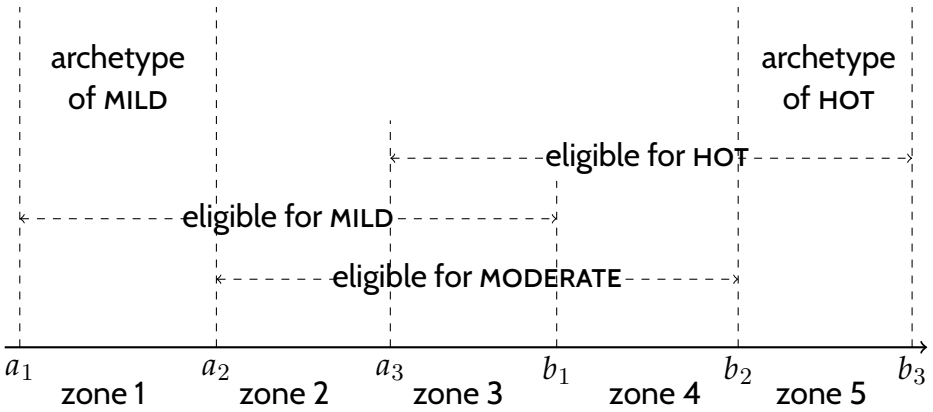

Figure 7.3: Three concepts: non-simple language

In figure 7.3, dishes in zone 3 (including $a_3 \& b_1$) are eligible for all three concepts, and no dish can be an archetype of MODERATE. This language is not simple because both boundaries are eligible to be placed anywhere from just below a_3 to just above b_1, including anywhere in zone 3.

In summary, the language is simple as it relates to a graded context just if no two boundaries can coincide and the archetype assumption holds.

7.4 Examples

The following two examples illustrate the parallels to theorems 1 and 2. In each example, there are exactly five dishes.

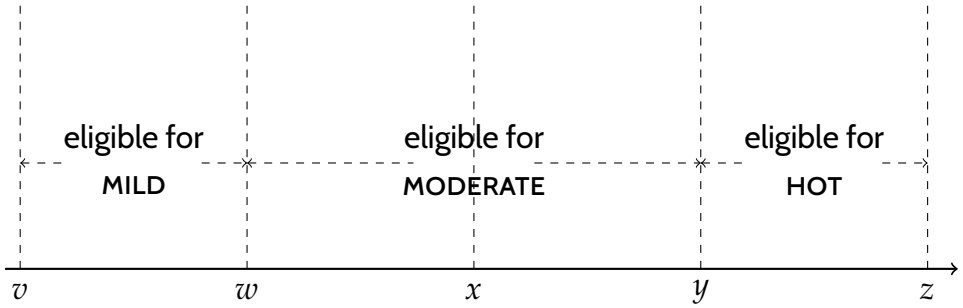

Figure 7.4: Three concepts: simple example

Figure 7.4 shows a simple language in the which the boundary between MILD and MODERATE is eligible to be placed between $v \& w$ and between $w \& x$, and the boundary between MODERATE and HOT is eligible to be placed between $x \& y$ and between $y \& z$. So v, x, z are archetypes and w, y are each eligible for two concepts. The four classifications that are compatible with figure 7.4 are given in table 7.1.

MILD	MODERATE	HOT
$\{v, w\}$	$\{x\}$	$\{y, z\}$
$\{v\}$	$\{w, x\}$	$\{y, z\}$
$\{v, w\}$	$\{x, y\}$	$\{z\}$
$\{v\}$	$\{w, x, y\}$	$\{z\}$

Table 7.1: Simple language: classifications

Rosa's influencers can each state one of these classifications, and Rosa can derive her two boundary locations by following a majority, considering each boundary independently. She can nominate a tie-breaking influencer if necessary - the same for both boundary placements.

The second example illustrates a non-simple language in which the boundary between MILD and MODERATE is eligible to be placed between $v\&w$, between $w\&x$ and between $x\&y$, and the boundary between MODERATE and HOT is eligible to be placed between $w\&x$, between $x\&y$ and between $y\&z$. So, x is eligible for all three concepts, and there is no archetype of MODERATE. There are two places for which both boundaries are eligible.

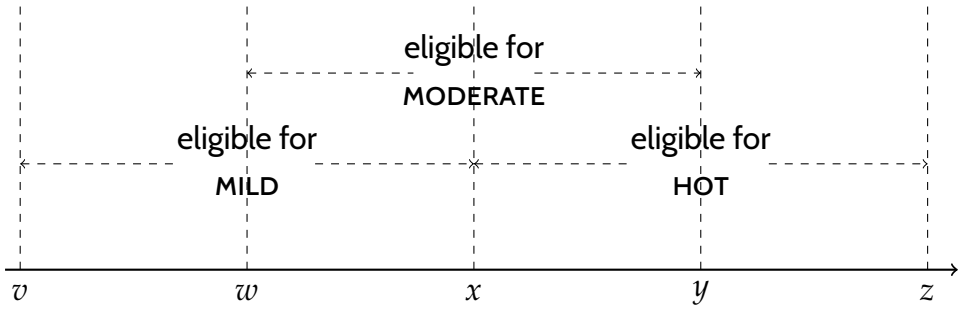

Figure 7.5: Three concepts: non-simple example

MILD	MODERATE	HOT
$\{v, w\}$	$\{x\}$	$\{y, z\}$
$\{v\}$	$\{w, x\}$	$\{y, z\}$
$\{v, w\}$	$\{x, y\}$	$\{z\}$
$\{v\}$	$\{w, x, y\}$	$\{z\}$
$\{v, w, x\}$	$\{y\}$	$\{z\}$
$\{v\}$	$\{w\}$	$\{x, y, z\}$

Table 7.2: Non-simple language: classifications

In this language, the influencers can state any of the classifications in table 7.2. Then there are some profiles of influencers' classifications for which Rosa has no majority outcome for the placing of one or other of the boundaries as illustrated by the profile of table 7.3.

influencers	MILD	MODERATE	HOT
1,2,3	$\{v, w, x\}$	$\{y\}$	$\{z\}$
4,5	$\{v, w\}$	$\{x\}$	$\{y, z\}$
6,7	$\{v\}$	$\{w, x\}$	$\{y, z\}$
Rosa	$\{v, w, x\}$	$\emptyset$	$\{y, z\}$

Table 7.3: Preponderance example from figure 7.5

Three influencers place the boundary between MILD & MODERATE between $x\&y$, and two place that boundary in each of its other eligible locations. A straightforward extension that Rosa might use is to place each boundary using a preponderance methodology, but this leads to possibility exemplified by the profile of table 7.3. The preponderant set of influencers place the boundary between MILD & MODERATE between $x\&y$, and a majority (also, of course, a preponderance) place the boundary between MODERATE & HOT between $x\&y$. So, Rosa places more boundaries between $x\&y$ than does any influencer. This violates a deference principle

based on unanimity (section 7.5) and entails that Rosa, unlike any of her influencers, leaves one concept empty

7.5 Deference principles

The principles used in considering deference in a graded context focus on placing boundaries between concepts rather than directly with the locations of objects, although, of course, these are intimately related.

Influencers can disagree about the location of a specified boundary (such as that between MILD & MODERATE). Also, when the language is not simple, they can agree that some boundary is placed between two dishes, but disagree about which one. Each of these is reflected in unanimity and consistency principles.

Boundary unanimity 1 Rosa does not place a specified boundary between two dishes $x\&y$ unless some influencer places that specified boundary between $x\&y$.

This principle entails that Rosa does not place a boundary where it is not eligible to be placed, so that Rosa follows her influencers in conforming to the eligibilities of the language. It also entails that Rosa places a specified boundary between $x\&y$ if every influencer places that specified boundary between $x\&y$, because no influencer places that boundary anywhere else.

Boundary unanimity 2 The number of boundaries placed by Rosa between $x\&y$ is not outside the range of the number of boundaries placed by any influencer between $x\&y$.

This principle applies when influencers place different boundaries between $x\&y$. It does not itself exclude the possibility that Rosa places the MILD-MODERATE bound-

ary between two dishes when every influencer places the MODERATE-HOT boundary between those two dishes. This possibility is excluded by ***boundary unanimity 1***. No influencer states an empty concept, entailing that no influencer places both boundaries between a pair of neighbouring dishes. ***Boundary unanimity 2*** entails that Rosa also follows this restriction.

The first consistency principle concerns a specified boundary.

> ***Boundary consistency 1*** In comparing two profiles, Rosa does not change whether she places a specified boundary above or below x unless some influencer changes whether he places that specified boundary above or below x.

The evidence required to establish whether Rosa places a specified boundary above or below a named dish is limited to whether her influencers place that boundary above or below that dish.

> ***Boundary consistency 2*** In comparing two profiles, Rosa does not change whether she places a boundary (not necessarily a specified boundary) just above x unless some influencer changes whether he places a boundary just above x.

The evidence required to establish whether Rosa places some boundary just above a dish is limited to whether her influencers locates a boundary there. If that evidence does not change, neither does' Rosa's conclusion on placing some boundary there. This principle is irrelevant in when the language is simple.

Consistency principles establish Rosa's reaction when no influencer changes some aspect of his classification. Monotonicity principles exclude perverse reactions when some influencer makes a change, and the proof of theorem 4 in chapter 11 requires such a deference principle to supplement ***boundary consistency 1***.

Boundary monotonicity In comparing two profiles, if additional influencers place a specified boundary just above x then Rosa does not move that specified boundary from just above x.

Finally, the analysis of deference in a graded context involves a symmetry principle that requires Rosa to follow a particular change by all of her influencers.

Boundary symmetry In comparing two profiles, if every influencer places a specified boundary just below or just above x, and then every influencer switches from placing that specified boundary just below x to placing it just above x, or *vice versa*, then Rosa also switches from placing that boundary just below x to placing it just above x, or *vice versa*.

So Rosa follows a unanimous move by her influencers that switches the location of a boundary from just above x to just below x or *vice versa*.

Each of the above principles reflects Rosa's decision to follow when 'every influencer states Φ', or 'no influencer states Φ'. Each (separately) is satisfied by a majoritarian and preponderance aggregators. For example, if Rosa places a specified boundary above x just if a majority do so, then ***boundary symmetry*** holds. Theorem 4 examines other aggregators, including those that do not rely on 'counting heads' to assess the weight of evidence for and against a specific outcome. The conditions apply only when the evidence is unanimously stated, and together they imply that Rosa must nominate a dominant influencer.

7.6 The theorems

In a graded context, theorems 3 and 4 are the parallels to theorems 1 and 2. They are proved formally in chapter 11.

Theorem 3. *If the language is simple, Rosa can place boundaries between non-empty concepts in a way that is responsive to each influencer and that satisfies* **boundary unanimity 1 & 2**, **boundary consistency 1 & 2**, **boundary monotonicity** *and* **boundary symmetry**.

This can be illustrated by the use of a majority aggregator as is possible for theorem 1. In a simple language,

Theorem 4. *If the language is not simple, Rosa can locate dishes into non-empty concepts in a way that satisfies* **boundary unanimity 1 & 2**, **boundary consistency 1 & 2**, **boundary monotonicity** *and* **boundary symmetry** *only by nominating a dominant influencer.*

Theorem 4 applies to the configuration of figure 7.5 on page 105. Under the conditions of the theorem, Rosa must nominate a dominant influencer. That dominance extends to configurations such as those of figures 7.6 and 7.7 even though in both figures z is an archetype of HOT and no dish is eligible for all of MODERATE, HOT and HOTTER.

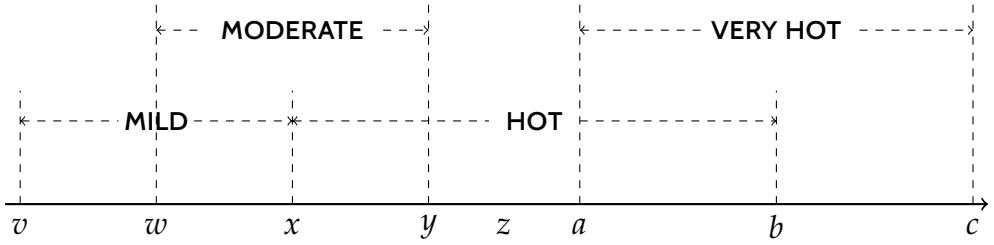

Figure 7.6: Four concepts: non-simple configuration

influencers	MILD	MODERATE	HOT	VERY HOT
1	$\{v, w\}$	$\{x, y\}$	$\{z, a\}$	$\{b, c\}$
all others	$\{v, w\}$	$\{x\}$	$\{y, z\}$	$\{a, b, c\}$
Rosa	$\{v, w\}$	$\{x, y\}$	$\{z\}$	$\{a, b, c\}$

Table 7.4: Profile for figure 7.6

If Rosa were concerned only with placing the boundaries between MILD & MODERATE and between MODERATE & HOT, theorem 4 entails that she must nominate a dominant influencer (say, influencer 1) if she follows the deference principles included in the statement of theorem 4. On the other hand, if she were concerned only with placing the boundaries between HOT & VERY HOT she could follow the majority. That combination would entail that with the profile of table 7.4, Rosa places the boundary between MODERATE & HOT between $y\&z$ and the boundary between HOT & VERY HOT between $z\&a$. So, Rosa places two boundaries between $y\&a$ but each influencer places one boundary between $y\&a$ violating ***boundary unanimity 2***.

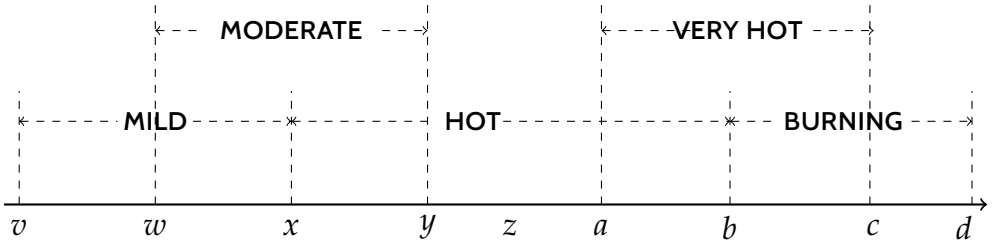

Figure 7.7: Five concepts: non-simple configuration

influencers	MILD	MODERATE	HOT	V. HOT	BURNING
1	$\{v, w\}$	$\{x\}$	$\{y, z\}$	$\{a, b, c\}$	$\{d\}$
all others	$\{v, w\}$	$\{x, y\}$	$\{z, a\}$	$\{b\}$	$\{c, d\}$
Rosa	$\{v, w\}$	$\{x\}$	$\{y, z, a\}$	$\{b\}$	$\{c, d\}$

Table 7.5: Profile for figure 7.7

A similar narrative applies to the configuration of figure 7.7 if Rosa nominates influencer 1 as dominant for the lower two boundaries and influencer 2 as dominant for the upper two boundaries. Then the profile of table 7.5 entails that Rosa places no boundary between $y\&a$ but every influencer places one boundary between $y\&a$ again violating ***boundary unanimity 2***.

7.7 Summary

Tis chapter introduces graded concepts, where the objects under consideration differ in only one relevant property. The illustrative context involves the property 'spiciness' and there is no disagreement about the relative spiciness of dishes. The concepts are ranked, and if x is spicier than y, then y cannot be located in a higher ranked category than x. This is a domain restriction that prevents the use of the analysis of chapter 4.

The remaining source of disagreement is the placing of boundaries between concepts, and the deference principles relate to the placing of boundaries rather than the location of objects. A simple language determines eligibilities that entail that no two boundaries are eligible to be in the same place. When the language is simple, there is scope for an archetype of each concept and Rosa can derive her concepts using a majority methodology.

When the language is not simple, there is no archetype of some concepts and deference principles entail that Rosa must nominate a dominant influencer. The deference principles involve unanimity and consistency principles in two forms related both to a specified boundary and to boundaries-in-general. There are, in addition, monotonicity and symmetry principles that relate only to the placement of specified boundaries. Proofs are given in chapter 11.

Chapter 8

Vagueness

8.1 Introduction

Multiple eligibilities - and hence the scope for individuals to make different statements within conformity to the language - can arise from vagueness of the concepts under consideration. For example, in the eponymous sorites narrative, a sufficiently large collection of grains of sand is eligible only for HEAP and a sufficiently small collection is eligible only for NOT-HEAP. In between there are collections that are eligible for both. Neighbouring collections differ by one grain and are indiscriminable, and the narrative concludes that there is no clear place for the boundary between HEAP and NOT-HEAP. When concepts are vague, individuals can appear to disagree because they state different precisifications of the concepts from those that are consistent with the eligibilities.

That differences result from vagueness adds little to the formal analysis of deference in chapters 4, 5 and 7. But when Rosa's influencers agree about eligibilities but state different precisifications, there is a sense in which they do not disagree fundamentally. They conform to the language, and in the simple case illustrated in figure 7.1 (page 102) they agree that dishes in zone 1 are in MILD, those in zone 3 are in HOT, and those in zone 2 could be in either concept, and no influencer has

any reason to privilege[34] one of the permitted precisifications over any other. If an influencer is asked 'what is the location of dish x?', they could give one of three answers: 'MILD', 'HOT' or 'could be either MILD or HOT'. Given these three possible answers, every influencer will give the same answer for x, depending on its zone in figure 7.1. A simple unanimity principle would entail that Rosa agrees and gives the same answer as all of her influencers[35].

A more subtle question arises if the influencers must answer 'MILD'or 'HOT', arbitrarily choosing one of these answers for dishes in zone 2. Section 8.2 shows that there is a sense in which Rosa does not fully reflect the answers given by her influencers.

A more significant difference from the previous analysis arises when the language itself is vague. In chapter 3 a different set of eligibilities entails different languages, and it was assumed that all of Rosa's influencers conform to the language, and so are constrained by the same eligibilities. If the language is vague the demarcation of eligibility sets is not clear, and each influencer might choose a different precisification of the language. In its simplest form, where Rosa considers only two concepts, MILD and HOT, the vagueness of the language is revealed in multiple possibilities for the three sets: archetypes of MILD, archetypes of HOT and archetypes of MILD. A similar distinction could arise if she is considering JAZZ and BLUES. This is an interpretation of second-order vagueness in the two basic concepts. For clarity, we refer to the precisification of the vague language that is used by an individual as that individual's **idiolect**

This leads to two further questions.

[34] A term borrowed from physical science where a privileged frame of reference has some special status or offers some simplification compared to others. In special relativity theory there are no privileged inertial frames.

[35] This is a different situation from that discussed in section 5.1 where an allowed answer is 'both', which is the default response by Rosa only when there is insufficient support for either 'HOT' or for 'MILD'.

- How does our terminology involving eligibilities and archetypes relate to that used routinely in discussions of vagueness, where objects are said to be possibly or definitely in the extensions of concepts? (section 8.3)

- If language is not precise, so that the eligibilities are themselves vague, does Rosa face additional challenges? (section 8.4)

8.2 The centralising tendency

If Rosa's influencers conform to a precise language, they agree on the eligibilities of dishes for the two concepts HOT and MILD, but may state different precisifications. The implications are exactly as in chapter 7: Rosa can follow the deference principles of theorem 3 if she places the boundary between HOT and MILD using a majority methodology.

If the differences between her influencers' classifications are attributed only to vagueness, there is no reason to suppose that any influencer privileges one precisification over any other. So, there is no reason to believe that they are more likely to use one precisification than any other. But following the majority entails that Rosa is more likely to use some classifications rather than others. There is a **centralising tendency**.

To illustrate this, suppose that there are four dishes v, w, x, y that are all eligible for both MILD and for HOT, plus archetypes of MILD (u) and of HOT (z). The grading relation gives $z \succ y \succ x \succ w \succ v \succ u$. The five possible precisifications are shown in table 8.1. Rosa follows three influencers and so there are 125 possible profiles of precisifications, one stated by each influencer. Following the majority leads Rosa to one of the five possible precisifications. The 'number' column shows how many of the 125 profiles result in each outcome.

outcome classification		number	frequency
MILD	HOT	of profiles	
$\{u, v, w, x, y\}$	$\{z\}$	13	10.4%
$\{u, v, w, x\}$	$\{y, z\}$	31	24.8%
$\{u, v, w\}$	$\{x, y, z\}$	37	29.6%
$\{u, v\}$	$\{w, x, y, z\}$	31	24.8%
$\{u\}$	$\{v, w, x, y, z\}$	13	10.4%
total		125	100%

Table 8.1: Centralising tendency: simple majority methodology

The 'central' classification $\{u, v, w\}, \{x, y, z\}$ results from following the majority for more profiles than any other - and the 'extreme' classification result from fewer. The boundary is below y in $^4/_5$ = 80% of the possible precisifications for any influencer, but is below y in $^{112}/_{125}$ = 89.6% of the classifications as they are devised by Rosa. Rosa locates x and y more frequently in HOT and v and w more frequently in MILD than is the case for influencers. There is some tendency for Rosa to privilege some precisifications even though no influencer privileges any precisification. In this case, Rosa's deference does not lead her to an exact replication of the unclarity that all of her influencers face.

This example raises the question (not answered here) of whether continued interaction between individuals, who might be each others' influencers, would reinforce the centralising tendency. Then an initial supposition that no precisification is privileged would be supplanted by a process that leads to the more frequent use of the central precisifications. In this scenario, mutual interaction might lead to some privileging of precisifications even though no individual has a personal reason (as opposed to their response to others), or any evidence, that supports that privileging.

8.2.1 Many-valued logic

Much of the literature on vagueness concerns the difficulty of establishing truth values of propositions. One account involves the use of many-valued logic and fractional 'degrees of truth' to circumvent the issues that arise in binary logic when there is vagueness[36]. The centralising tendency can entail that, if Rosa defers to the majority view of her influencers, she does not reflect unanimously held fractional truth values.

Suppose that the truth value of a proposition is defined to be the proportion of all possible precisifications in which that proposition is true[37]. Then in the example of table 12.1, every influencer assigns a fractional truth value of 0.8 to the statement "y is in HOT" ≡ "the boundary is below y" because the statement is true in $4/5$ of the possible precisifications. If Rosa aggregates her influencers' classifications by following the majority, her truth value of "y is in HOT" is 0.896 ($= {}^{112}/_{125}$). The aggregation methodology entails that Rosa's fractional truth value does not reflect that unanimously held by her influencers. So this interpretation of the truth values of vague propositions fails to satisfy a simple unanimity principle when classifications are aggregated.

8.3 Terminology and taxonomy

Eligibilities are defined by the language of the community, but the literature on vagueness does not refer to eligibilities. Instead, using the illustrative context of chapter 7 with two concepts, dishes that are archetypes of (eligible only for) HOT are termed **definitely HOT**, those that archetypes of MILD are **definitely MILD** (or definitely NOT-HOT), and those that are eligible for both are **possibly HOT** (and possibly MILD). Every statement that we make about archetypes can be transformed into statements involving 'definitely', and statements about eligibilities can

[36] Chapter 4 of Williamson (1994) examines issues that arise in this approach.

[37] A parallel result can be derived as long as the degree of truth is an increasing function of the proportion of all possible precisifications in which that proposition is true.

be transformed into statements involving 'possibly.' The converse transformation of terminology is also possible.

We have used 'eligible for' to allow that there is a truth of the matter but that is unknowable (epistemic interpretations of vagueness) or not yet known (medical diagnoses in the current state of technology), and to allow that apparently contradictory statements are relative to the individual making them (a personal evaluation of the properties and history of a composition). Suitably read, 'it is possibly true that my knee-pain exemplifies LUPUS' might be regarded as true, but is in fact false, as revealed by a later (perhaps *post mortem*) diagnosis. Saying instead 'my knee pain is eligible for LUPUS' arguably avoids the question of whether a statement that is actually false can be possibly true. The doubt entailed by saying '*Rhapsody in Blue* is possibly CLASSICAL and possibly JAZZ' does not reflect the position of a musicologist who locates *Rhapsody* in CLASSICAL following debate with others and consideration of the properties and history of the composition. The musicologist can allow that others might take a different view (which he does not regard as true) by locating *Rhapsody* in JAZZ but cannot allow that anyone can locate *Rhapsody* in OPERA. In both these examples, the use of 'eligibility' arguably avoids some questions that surround the use of 'possibly true'. But this informal commentary does not prevent formal the restatement of our earlier analysis in an alternative form.

8.3.1 Second-order vagueness

First-order vagueness in MILD and HOT arises when there are clear boundaries between three sets labelled *archetypes of HOT, archetypes of neither* and *archetypes of MILD* (the zones in figure 7.1).

Second-order vagueness in MILD and HOT arises when there is unclarity about the location of the boundaries between these three sets of dishes. Eligibilities, and hence the language itself are vague. Figures 8.1 and 8.2 show the two possibilities for the taxonomy of second-order vagueness that parallel figures 7.2 and 7.3.

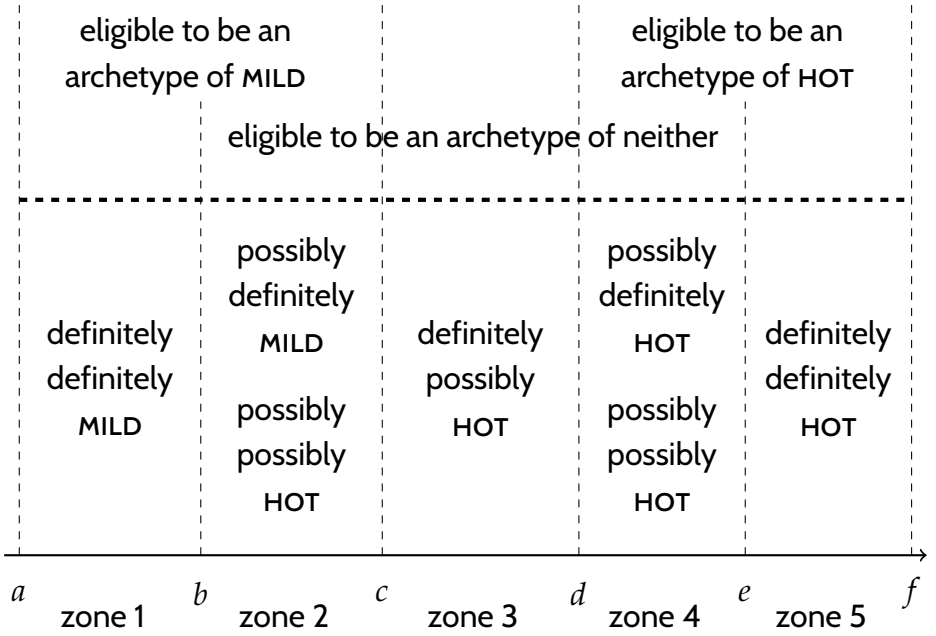

Figure 8.1: Taxonomy of second-order vagueness: configuration A

Figure 8.1 reproduces figure 7.4 except that the three concepts MILD, MODERATE and HOT have been replaced respectively by the three sets of dishes. The boundary between dishes that are archetypes of MILD and those that are archetypes of neither is just below b, in zone 2, or just above c. The boundary between dishes that are archetypes of neither and those that are archetypes of HOT is just below d, in zone 4, or just above e.

Figure 8.1 also shows the labelling of the five zones using 'definitely' and 'possibly', recognising that with two ranked concepts 'possibly HOT' ≡ 'possibly MILD'.

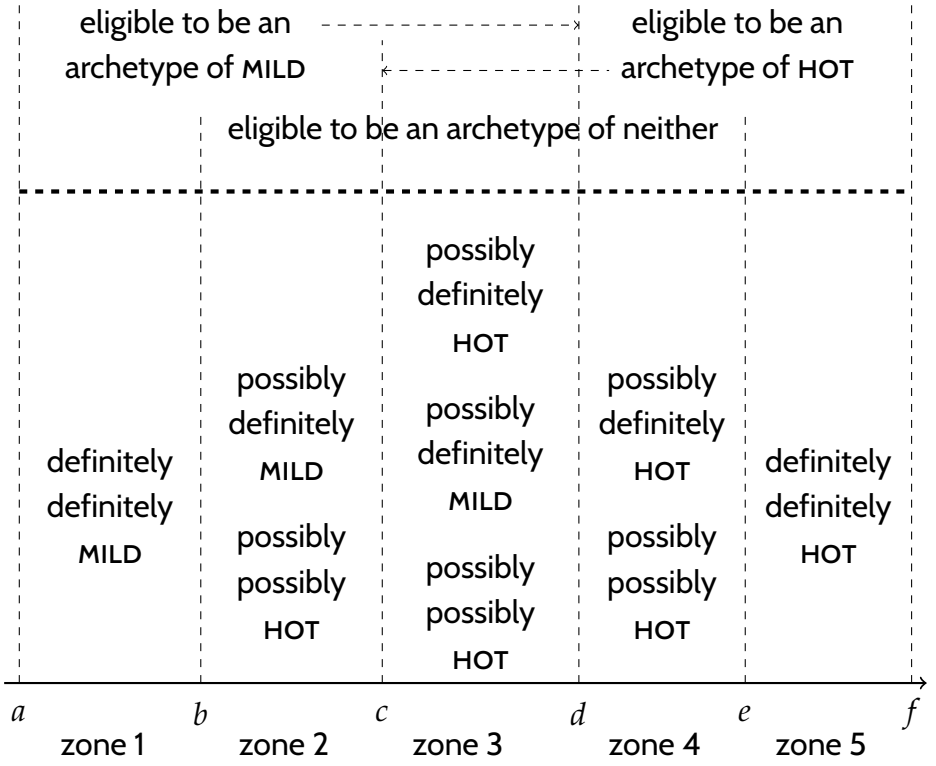

Figure 8.2: Taxonomy of second-order vagueness: configuration B

Figure 8.2 similarly reproduces figure 7.5. The boundary between dishes that are archetypes of MILD and those that are archetypes of neither is just below b, in zones 2 or 3, or just above d. The boundary between dishes that are archetypes of neither and those that are archetypes of HOT is just below c, in zones 3 or 4, or just above e. Figures 8.1 and 8.2 differ in both formulations in their descriptions of zone 3.

8.3.2 Williamson's taxonomy

Figures 8.1 and 8.2 show alternative configurations in the taxonomy of second-order vagueness. They refine the taxonomy given as the opening statement of Williamson (1999) as it applies to graded objects and ranked concepts.

Williamson's taxonomy, applied to the vague concept BALD is

> People can be bald or not bald. That classification is vague. People can be definitely bald or definitely not bald, but they can also be nei-

> ther definitely bald nor definitely not bald. That classification is vague too. People can be definitely definitely bald or definitely definitely not bald or definitely neither definitely bald or definitely not bald, but they can also be neither definitely definitely bald nor definitely not definitely bald, or neither definitely definitely not bald nor definitely not definitely not bald or neither definitely neither definitely bald nor definitely not bald nor definitely not neither definitely bald nor definitely not bald.[38]

If heads are graded according to their numbers of hairs[39], and we substitute 'mild' for 'bald', 'hot' for 'not-bald', and 'dishes' for 'people', we have the following from the Williamson quote:

- Dishes can be mild or hot.

  > That classification is vague, within precise eligibilities that allow first-order vagueness in MILD and HOT.

- First-order vagueness in MILD and HOT entails that dishes can be definitely mild [≡ an archetype of MILD] or definitely hot [≡ an archetype of HOT], but they can also be neither definitely mild nor definitely hot [≡ possibly hot ≡ an archetype of neither].

  > That classification can also be vague, allowing vagueness in the eligibilities defined by the language, which is second-order vagueness in MILD and HOT.

- Second-order vagueness in MILD and HOT entails that a dish can be in one of six categories: W(i) to W(vi):

[38] Writing such a statement if there are three ranked concepts would be a formidable task using only 'definitely' and 'not'. Using eligibilities would be simpler.

[39] Williamson does not specify this, but much writing on vagueness assumes a graded context, including the original sorites

W(i) definitely definitely mild [≡ eligible only to be an archetype of MILD] - in zone 1 in figures 8.1 and 8.2.

W(ii) definitely definitely hot [≡ eligible only to be an archetype of HOT] - in zone 5 in figures 8.1 and 8.2.

W(iii) definitely neither definitely mild nor definitely hot ≡ definitely possibly mild [≡ eligible only to be an archetype of neither] - in zone 3 in figure 8.1; not present in figure 8.2.

W(iv) neither definitely definitely mild nor definitely not definitely mild ≡ possibly definitely mild [≡ eligible to be an archetype of MILD and an archetype of neither] - in zone 2 in figure 8.1, zones 2 and 3 in figure 8.2.

W(v) neither definitely definitely hot nor definitely not definitely hot ≡ possibly definitely hot ≡ possibly definitely mild [≡ eligible to be an archetype of HOT and an archetype of neither] - in zone 4 in figure 8.1, zones 3 and 4 in figure 8.2.

W(vi) neither definitely neither definitely mild nor definitely hot nor definitely not neither definitely mild nor definitely hot ≡ neither definitely possibly mild nor definitely not possibly mild ≡ possibly possibly mild [eligible to be an archetype of MILD, an archetype of HOT, an archetype of neither] - not present in figure 8.1; zone 3 in figure 8.2.

The equivalence is given in table 8.2.

zone	figure 8.1	figure 8.2
1	W(i)	W(i)
2	W(iv)	in W(iv), not in W(vi)
3	W(iii)	W(vi)
4	W(v)	in W(v), not in W(vi)
5	W(ii)	W(ii)
absent	W(vi)	W(iii)

Table 8.2: Taxonomy of second-order vagueness

Of Williamson's six categories only five can co-exist in a graded context. The identity of the absent category depends on the language.

8.4 Vagueness of the language

When the language is vague, each of Rosa's influencers might have a different idiolect which is a single precisification from the several that the vague language allows. Effectively, in a two-concept, graded context, an idiolect defines the placement by an individual of two boundaries between three categories of dishes: archetypes of MILD, archetypes of neither concept, archetypes of HOT. In obvious terminology, these are the **lower boundary** and the **upper boundary**.

So Rosa faces a challenge in two stages: she can define her own idiolect through deference to the idiolects stated by her influencers, and she can define her classification within her idiolect through deference to the boundaries stated by her influencers. In the first stage, she defers to her influencers to determine her own idiolect in the form of the three sets: the archetypes of each concept and archetypes of neither. At the second stage, she defers to form her two concepts. We can presume that Rosa wants to the two stages to yield compatible outcomes in the sense that if, at the first stage, she identifies dish x as an archetype of one concept, she does not locate x in a different concept at the second stage.

8.4.1 A compatible configuration

Given earlier conclusions, it is perhaps unsurprising that the possibility of incompatibility between the two stages depends on the configuration of eligibilities for the three categories.

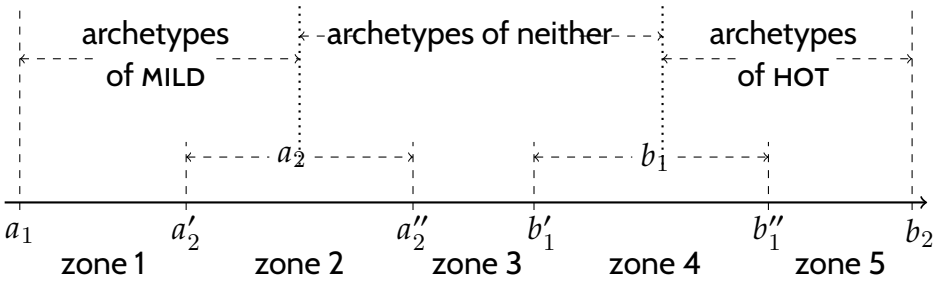

Figure 8.3: Vagueness of eligibilities: configuration A

Each influencer places the lower boundary just below a_2', in zone 2 or just above a_2''. Likewise, each influencer places the upper boundary just below b_1', in zone 4 or just above b_1''.

Rosa's idiolect is determined by her determination of the positions of a_2 and b_1, and we suppose that she derives this by deferring to her influencers. In figure 8.3, these can be derived from the majority view of her influencers, so that Rosa regards a dish in zone 2 as an archetype of MILD just if that is the stated view of a majority of her influencers.

As well as determining her idiolect, Rosa can place the boundary between MILD & HOT according to the majority view of her influencers. This dual process creates no incompatibility because any dish that Rosa regards as an archetype of HOT is regarded as an an archetype of HOT by a majority of her influencers, and so is located in HOT by a majority of influencers, and hence by Rosa.

8.4.2 Potential incompatibility

Figure 8.4 shows the alternative configuration in which dishes in zone 3 are eligible to be archetypes of MILD, and of HOT, and of neither. Three possibilities are shown in the figure with examples of locations for a_2 and b_1. Each influencer places the lower boundary just below a_2', in zone 2 or zone 3 or just above a_2''. Likewise, each influencer places the upper boundary just below b_1', in zone 3 or zone 4 or just above b_1''.

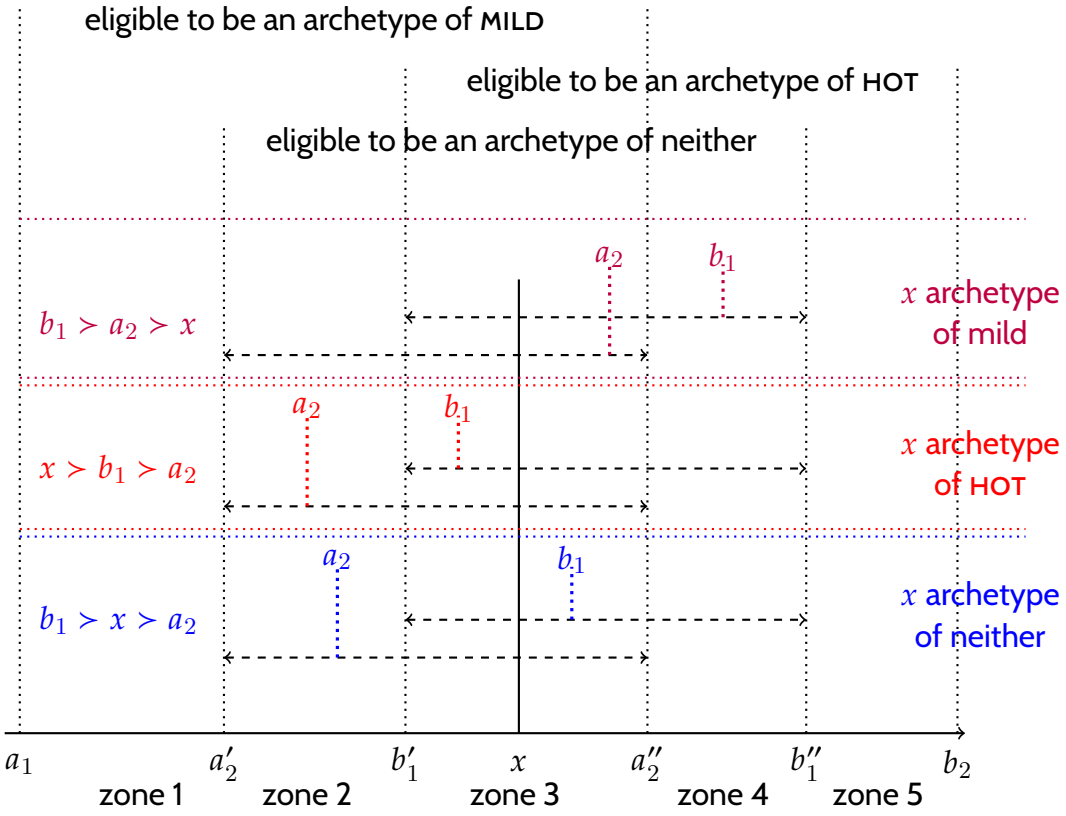

Figure 8.4: Vagueness of eligibilities: configuration B

If Rosa applies the deference principles of theorem 4 to her decision about the idiolect that she uses, Rosa must adopt the idiolect of a dominant influencer regardless of the idiolects of all other influencers. The scope for disagreement about the range of eligible idiolects is sufficient to entail that Rosa cannot compromise between the views of her influencers.

Having nominated a dominant influencer to determine her idiolect, Rosa can also determine the boundary between MILD & HOT. In principle she can do this by following the majority statement of her influencers. But this can mean that she

- follows her dominant influencer in, say, locating x in figure 8.4 as an archetype of MILD (the purple possibility with x below a_2), but
- follows the majority of her influencers who locate x in HOT (the red or blue possibility).

Rosa concludes that dish x is an archetype of MILD but also that x is in HOT. This incompatibility arises if the dominant influencer's idiolect is purple but a minority of influencers' idiolects are purple. With the extent of potential disagreement about idiolects that is typified in figure 8.4, Rosa can be certain to avoid incompatibility only if she nominates the same dominant influencer to determine her idiolect and to determine the placing of the boundary between MILD & HOT.

8.4.3 Higher-order vagueness in the language

First-order vagueness in the two concepts refers to the possibility that there are several possible places for the boundary between MILD & HOT. Vagueness of the language as it applies to the eligibilities of dishes in Rosa's set refers to the unclarity in the placing of the boundaries between archetypes of MILD and dishes that are archetypes of neither, and between those dishes and archetypes of HOT.

In considering the extent of vagueness of the language in figures 8.1 and 8.2, the positions of a_2 and b_1 are constrained between $a_2'\&a_2''$ and $b_1'\&b_1''$ respectively. It is possible that the positions of a_2', a_2'', b_1' and b_1'' are also unclear. There is then unclarity about the vague language within which Rosa and her influencers establish their idiolects. Each influencer must nominate a vague language, an idiolect and a location for each dish. This is a potential further source of incompatibilities for Rosa if she defers to her influencers using deference principles related to those used here, but the force of the argument is strongest when considering the

first- and second-order questions that give potential incompatibilities between her responses to vagueness of the language and to vagueness within the language. Arguably Rosa can articulate the challenges raised by these two orders of vagueness in MILD and HOT more easily than those raised by higher-order possibilities. However many orders of vagueness are considered, the conclusion remains that Rosa might need to nominate a dominant influencer even in establishing the single boundary between MILD & HOT if she is to avoid inconsistencies when she defers to influencers in devising her idiolect, and in forming her concepts within that idiolect.

8.5 Summary

This chapter extends the analysis to include the additional considerations raised by vagueness. The first is that, even if Rosa's influencers agree about the extent of vagueness in the formation of concepts, they can make different statements because they use different precisifications. Then Rosa's formation of concepts does not fully reflect that of her influencers who have no reason to privilege any precisification. Rosa, on the other hand, faces a centralising tendency that entails that she uses some precisifications in more circumstances than others.

A more fundamental consideration relates to second-order vagueness in the concepts. This can be interpreted as vagueness about eligibilities, and hence about the language itself. Rosa's influencers can then have different idiolects, which are precisifications of the vague language.

The distinction made in chapter 7 between simple and non-simple languages then becomes relevant. In a graded context, the analogue to a simple language has a different taxonomy from that of the analogue to a non-simple language. When there are two spiciness concepts, a simple language entails that no dish is eligible to be an archetype of MILD and an archetype of HOT, and some dishes are eligible to be an archetype of neither. A non-simple language entails that some dishes are

eligible to be an archetype of MILD and an archetype of HOT, and there is no dish that is eligible only to be an archetype of neither. In both cases, the taxonomy is more restricted than that of Williamson when the latter is applied in a graded context.

If Rosa first derives her own idiolect by deference to her influencers, and defers again to derive her concepts, the analogue to a simple language ensures that these two derivations are compatible. But with the alternative configuration, it is possible that Rosa can locate dish x in a concept at the second stage, having concluded that x is an archetype of a different concept at the first stage. She does this because at the first stage she defers to a dominant influencer, but follows the majority at the second stage. The majority might not include her first-stage dominant influencer.

Chapter 9

Designation

9.1 Introduction

This chapter concerns the formation of concepts whose extensions consist of members of the language community who are themselves using the concepts. For example, it could cover concepts such as JAZZ-MUSICIAN, BLUES-MUSICIAN ..., as they apply to the community, in which each individual can designate himself and others into one of these concepts. We retain a model in which Rosa forms her own concepts by considering those stated by other members of the language community[40]. This chapter assumes that the language entails that every individual is eligible for every concept. The exclusion by Rosa of an individual from a concept can follow the evidence from self- and other-designations, but is not determined *a priori*.

This example might also have a practical implication in allocating those who attend a convention of musicians who are to be partitioned into groups for performance or discussion of the genres. Each attendee designates himself and everyone

[40] She uses the opinions of all community members and does not defer to nominated influencers within the community. This allows the possibility that Rosa herself is a member of the community - but if she is, she reaches conclusions about her own location in a way that is not privileged compared with how she reaches conclusions about others.

else, and the conference organiser use these **self-** and **other-designations** to partition the convention into subgroups.

An alternative, and perhaps much more sensitive, context is the derivation of demographic concepts, for example by gender, by religion or by sexuality[41]. Rosa is then deciding who to include in her concept MALE or CHRISTIAN on the basis of community members' self- and other-designations.

In practical applications, inclusion in demographic categories can allow or restrict access to particular facilities, such as religious premises and gender-specific changing rooms, or affect participation in sports and other activities. There are many people who believe that an individual has the right to decide their category without intervention from others, and that non-discrimination principles then influence the consequences. Others disagree, arguing that not all self-designations should be allowed to pass unchallenged. This challenge leads people to reject the conclusions of national censuses or large social surveys because the only information collected is individual self-designations. Then, the argument proceeds, the census does not give acceptable extensions or anonymised counts of MALE or CHRISTIAN as these concepts apply to the community.

Kasher & Rubinstein (1997) formalise a context which arises from the Israeli 1950 *Law of Return* that states criteria for eligibility as a citizen (for example, The Jewish Agency, 2019, see also Kasher, 1985, 1993). These criteria are more extensive than the definition of Jewishness under Halachic law, and the resulting immigration has led to political controversy around, for example, the observance of the Sabbath (Reeves 2011, Galili 2020). In this context, the controversy involves whether the outcome for any individual should be based only on that individual's self-designation, or whether the 'weight of opinion' should allow other-designations to determine some outcomes. Kasher & Rubinstein formalise this problem as a vari-

[41]Or age, in which people use concepts and designate themselves and one another as OLD, MIDDLE-AGED.... This involves designation of a set of graded objects, but that variation is not pursued here.

ation in social choice theory. If Rosa takes into account potentially diverse self- and other-designations in forming her view of POTENTIAL ISRAELI CITIZEN, she needs some form of aggregator to reach her conclusion.

Social choice problems conventionally involve the aggregation of individual opinions concerning sets of objects. The unconventional aspect of the formulation by Kasher & Rubinstein is that the set of objects is identical to the set of individuals. Those to whom Rosa defers express self- and other-designations of themselves and each other. This need not affect the results of chapter 4 if Rosa uses ***class unanimity, pair unanimity*** and ***class consistency*** when aggregating these designations to form her own concepts[42]. With three or more concepts, theorem 2 entails that Rosa would need to follow the self- and other- designations of a dominant influencer.

Arguably - at least to liberals - an individual's self-designation has a different status from other-designations about or by that individual. An obvious alternative methodology is that Rosa locates each individual where that individual self-designates. This methodology clearly satisfies ***class unanimity*** and ***class consistency***, but there are profiles of designations for which Rosa's outcome does not satisfy ***pair unanimity***. Table 9.1 gives an example with four individuals and three concepts.

	designation of			
designation by	1	2	3	4
1	C_1	C_2	C_2	C_3
2	C_2	C_1	C_3	C_2
3	C_3	C_2	C_1	C_3
4	C_2	C_3	C_2	C_1

*Table 9.1: Self designation and **pair unanimity***

[42]Theorem 2 in Kasher & Rubinstein reaches this conclusion replacing ***pair unanimity*** with a requirement that Rosa leaves no concept empty: see also Cho & Ju (2017). The formal link between their approach and ours is explored in section 12.5.

In this example, a liberal Rosa co-locates all of the individuals in concept C_1 (using the red cells in the table) where they all self-designate, but no-one co-locates individuals 1&2 or 2&3 or 3&4.

Kasher & Rubinstein refine this analysis to show that there are combinations of principles that Rosa could follow that imply that she has no alternative to locating each individual where that individual self-designates. If she follows those principles, she is 'forced to be liberal'. However, Rosa need not reflect every self-designation if she follows only follows ***class unanimity*** and ***class consistency***. In the example of table 9.1 she does not violate these principles if she locates individual 1 where 2 designates 1, 2 where 3 designates 2, 3 where 4 designates 3 and 4 where 1 designates 4. In the example, Rosa would use the green cells to establish $C_1 = \emptyset, C_2 = \{1, 2, 3\}, C_3 = \{4\}$ as her concepts.

The remainder of this chapter refines the analysis of Kasher & Rubinstein by weakening some of their principles and extending from their limit of two concepts (an individual is either qualified to be an Israeli citizen or is not) to allow any finite number. The principles focus on the potential difference of status between self-designation and other-designation. A conclusion that Rosa cannot be other than liberal might be more palatable than the conclusion that she must nominate a dominant individual, but it still reflects that Rosa cannot compromise between the views of different individuals.

This unconventional extension of social choice does not lose the typical assumption that the only information that is relevant to the determination of the outcome of the social choice process are the individual opinions. If Rosa wishes to assert a concept that allows that an individual can be considered to be an Israeli citizen just if their mother has demonstrable Jewish ancestry, her information base extends beyond (and maybe excludes) the self- and other-designations invoked by Kasher & Rubinstein. Of course, any individual under consideration might base their self- and other-designations on that maternal principle, but the social choice approach allows (at least *ex ante*) that this opinion might not prevail, given the opinions of

others. The maternal principle is neither necessary nor sufficient to determine the outcome.

The demographic contexts are potentially controversial because there is a conflict between those who believe that self-designation is a liberal principle that ought to be respected by Rosa in her derived concepts, or by society at large in a practical application) and those who, at least in some cases, disagree. A formally similar problem arises in a context where (arguably) self-designation is unlikely to be supported as a general principle. Suppose that Rosa is a tutor who grades a group of students who have worked together on a project based on their assessments of their own and others' contributions. It might be felt to be reasonable for Rosa to take some note of an individual's self-assessment in determining her grade, but not to allow self-assessment to be the sole or dominant evidence used[43].

Much of the literature following Kasher & Rubinstein has been concerned with decision-making issues. For example, Samet & Schmeidler (2003) consider the extent of consent from others that is needed to allow an individual to self-designate, Alcantud & Laruelle (2020) consider the formation of 'clubs' using voting rules and involving vetoes, and Houy (2007) and Fioravanti & Tohmé (2020) constrain the influence of individuals so that, for example, existing members of a category have the privileged ability to admit others to that category. These are not our question, which remains the examination of the consequences of accepting that Rosa's derived concepts result only from self-designation by each individual, or of maintaining some role for other-designation.

[43]It is perhaps stretching a point to say that this methodology determines Rosa's understanding of concepts such as PASS, MERIT or FAIL, although she is willing to use the concepts in this way at least for this project assignment. Her comparison between these uses of the concepts and uses that arise elsewhere is not part of the consideration here.

9.2 Four principles

As previous chapters indicate, the social choice approach to the analysis of the consequences of disagreement puts forward principles that govern the way in which the individual statements are combined. These restrict the ways in which Rosa can respond to the diverse information that she faces. Kasher & Rubinstein follow this route, restricted to a division of the community into only two groups, and introduce a total of five principles (axioms in their terminology) on the aggregator that Rosa uses to combine the self- and other-designations. Variations on the first four principles are discussed in this section.

9.2.1 Unanimity

The first principle is straightforward and echoes those used previously:

> ***Designation unanimity***: if every individual (including i) designates i in concept C, then Rosa locates i in C.

Kasher & Rubinstein name this principle 'consensus'.

9.2.2 Consistency

Our previous discussion indicates[44] that if Rosa also uses a consistency (or independence) principle and both unanimity principles, she has no alternative to following the designation of a single dominant individual. Rosa would then use that individual's designation of himself and of everyone else in the community, and ignore all other designations. The possible special status of self-designation would then play no role in influencing Rosa's derivation of the concepts: she would agree with the self-designation of the dominant individual, but only with other self-designations if they were agreed by the dominant individual.

[44] Theorem 2 in Kasher & Rubinstein effectively confirms this suggestion: see also Cho & Ju (2017).

A standard independence principle requires that Rosa's location of i depends only on the individual's designations of i. To avoid the inevitability of the dominant individual, Kasher & Rubinstein considerably weaken the independence principle to require only that if one or more individuals change their designation of i, and no other designation is changed, it is not the case that the only impact on Rosa is that she changes her location of some other individual $j \neq i$.

> ***Designation consistency***: If no individual (including j) changes his designation of j, and some individual changes some self- or other- designation of $i \neq j$, then Rosa changes her location of j only if she also changes her location of at least one other individual (who need not be i).

By adhering to ***designation consistency***, Rosa rules out the use of aggregators that might technically satisfy other principles, but that seem to have no underlying justification. An example of such an aggregator would arise if Rosa were to arrange the individuals in a circle, and locate each individual in the concept to which that individual designates his left-hand neighbour[45]. If j is the left-hand neighbour of i, and i designates j in C then Rosa locates i in C. A change in Rosa's location of i is the only result of change in the other-designation of j by i. This violates ***designation consistency***.

Designation consistency has analogues in other interconnected contexts such as multi-market analysis. For example, if there is a climate-related reduction in the production of coffee beans, then a partial equilibrium analysis of the coffee market alone asserts that the price of coffee (as paid by the coffee-drinker) rises. A general market analysis indicates that people switch to other drinks so that the price of cocoa rises, and also that the price of coffee-makers is likely to fall because fewer are

[45] Perhaps this is a retaliatory methodology: i is located where i designates his neighbour. It is also of course easily manipulable if the individuals understand and care about what Rosa is doing. In other areas of social choice theory, manipulability is closely related to the violation of an independence condition (see Craven, 1992).

used. Further income effects on many other markets might follow because people who drink a lot of coffee have less to spend on other things. So, the price of a good can change either because of direct effects of incidents in its own market, or because of indirect effects from incidents in other markets. What is extremely unlikely (excluded by an analogue of ***designation consistency***) is that the only impact of the reduction in the availability of coffee is a change in the price of cocoa.

Designation consistency allows these sorts of interactions or **system effects** in the analysis of Rosa's derivation of concepts. For example, suppose that some individual re-designates j, and that this leads to a change in Rosa's location of j. Many other individuals might feel that k should be co-located with j because of their similarities, and an aggregator that reflects this widespread view of co-location would also change the outcome location of k, even though no individual has re-designated k. ***Designation consistency*** allows these additional effects: a stronger principle could rule them out, but that is not needed here.

9.2.3 Monotonicity

In conventional social choice, an independence condition is often accompanied by a monotonicity condition, for the simple reason that the former specifies that additional support for some outcome can change that outcome, but does not rule out the possibility that the change is perverse. A monotonicity condition ensures that additional support reinforces an outcome decision. The independence condition here is ***designation consistency***, and a monotonicity condition can be most clearly stated when system effects do not occur.

> ***Designation monotonicity***: If the only change in designations is that some individuals re-designate j into C and if the consequence of the re-designations is that Rosa does not re-locate any individual other than j, then Rosa does not re-locate j out of C.

One circumstance in which this principle can be applied arises if Rosa follows ***designation unanimity*** and if every individual other than j is unanimously designated in a concept. Then if ***designation monotonicity*** holds, and additional individuals re-designate j into C, Rosa does not perversely re-locate j out of C.

9.2.4 Symmetry and non-discrimination

In conventional social choice an anonymity condition ensures that, if two individuals' opinions about the objects are swapped, then the outcomes from the profiles are the same. A symmetry condition ensures that if two objects are swapped in every individual's opinion, then the outcomes for the two objects are also swapped. Anonymity can cause difficulties when 'head-counting' aggregators such as that based on the simple majority methodology are used because some individual must be nominated as tie-breaker.

In the designation context, a symmetry condition would require Rosa to co-locate $i\&j$ if their positions are swapped both in their role as designators, and in their role as designatees. Cho & Ju (2017) have asserted that 'name changes shift no fundamental content' (page 519) and Samet & Schmeidler (2003) that 'a rule does not depend on the names of individuals' (page 216). They invoke symmetry conditions that have relatively wide applicability.

A principle that generalises that used for two categories in Kasher & Rubinstein would be:

Designation symmetry: Rosa co-locates individuals i and j if all of the following are true:

(i) i self-designates in the same concept that j self-designates;

(ii) each other individual k designates i in the same concept that k designates j;

(iii) i designates j in the same concept that j designates i;

(iv) i designates each other individual k in the same concept that j designates k.

Part (iv) of ***designation symmetry*** is arguably less attractive than the other parts. It seems reasonable to suppose that Rosa co-locates two individuals i and j who are similarly self- and other-designated because she has no information that she can use to discriminate between them (parts (i), (ii) & (iii)). But it seems less attractive to base this non-discrimination in her treatment of i and j on the additional requirement that i and j designate other individuals in the same way.

One circumstance in which parts (i) to (iii) but not part (iv) of ***designation symmetry*** can occur is if two individuals are both $s\backslash t$ **singletons**. Individual i is identified as an $s\backslash t$ singleton when:

(i) i self-designates in C_s;

(ii) every other individual designates i in C_t.

If i and j are both $s\backslash t$ singletons, then both i designates j and j designates i in C_t, satisfying part (iii) of ***designation symmetry***. Trivially parts (i) and (ii) are also satisfied. Part (iv) would not be satisfied if the $s\backslash t$ singletons did not agree on the location of some other individual.

A weaker symmetry principle with less extensive applicability than ***designation symmetry*** is:

Non-discrimination: Rosa co-locates all $s \setminus t$ singletons.

Unlike ***designation symmetry***, this principle does not require that Rosa takes into account the designations made by the $s \setminus t$ singletons: it does not involve part (iv) of ***designation symmetry***.

9.3 Majority and preponderance

Kasher & Rubinstein assume that the domain of the aggregator is not restricted because every individual is eligible for inclusion in both concepts[46]. But they cannot rely on majority methodologies because their fifth principle prevents its use (section 9.4). The other four principles (***designation unanimity***, ***designation independence***, ***designation monotonicity***, ***designation symmetry***) permit the use of a majority methodology when only two concepts are considered. Rosa can resolve ties by locating an individual in the concept in which he self-designates. Substituting ***non-discrimination*** for ***designation symmetry*** does not change this conclusion.

Maintaining the assumption that every individual is eligible for every concept[47] entails that there might not be a majority designation when there are three or more concepts. A preponderance aggregator locates i in the concept that is supported by the largest number of individuals (including i), but this needs a tie-breaking mechanism, which now requires that Rosa establishes a hierarchy of tie-breakers because ties can arise in several ways.

[46] In their context, no individual is *a priori* excluded from or automatically included in Israeli citizenship.

[47] Violating this assumption involves the application of an *a priori* eligibility restriction that sits uncomfortably with the 'liberal' view that each individual can self- and other-designate freely.

	designation of individual 1	
designation by	Profile 9.2a	Profile 9.2b
1	C_1	C_1
2	C_1	C_2
3	C_2	C_2
4	C_2	C_3
5	C_3	C_3

Table 9.2: Preponderance designations

If Rosa nominates individual 1 as the first level tie-breaker for himself, then for profiles consistent with 9.2a in table 9.2, she locates individual 1 in C_1. For profiles consistent with 9.2b, there is no tie involving individual 1, and Rosa needs to nominate an alternative, second-level, tie-breaker in order to locate individual 1 either in C_2 or in C_3. This preponderance aggregator satisfies ***designation unanimity***, ***designation independence*** and ***designation monotonicity***. It also satisfies ***non-discrimination*** because the preponderance methodology does not locate any singleton where she self-designates (assuming that there are more than two individuals), and so Rosa locates all $s \backslash t$ singletons in C_t.

However, this methodology does not satisfy ***designation symmetry*** because of the 'privileged position' of the second-level tie-breaker. For example, suppose that Rosa nominates:

- each individual as his own first-level tie-breaker;
- 3 as the second-level tie-breaker for locating 1;
- 4 as the second-level tie-breaker for locating 3.

The designations of table 9.3 entail that Rosa locates individual 1 in C_2 and individual 3 in C_3.

	designation of				
designation by	1	2	3	4	5
1	C_1	C_2	C_2	C_3	C_3
2	C_2	C_2	C_2	C_3	C_3
3	C_2	C_2	C_1	C_3	C_3
4	C_3	C_2	C_3	C_3	C_3
5	C_3	C_2	C_3	C_3	C_3
tie-breaker	1 then 3		3 then 4		
location by Rosa	C_2	C_2	C_3	C_3	C_3

Table 9.3: Example with two levels of tie-breaker

With the profile of table 9.3

(i) 1 and 3 both self-designate in C_1.

(ii) each other individual designates 1 and 3 into the same category.

(iii) 1 and 3 mutually other-designate in C_2.

(iv) 1 and 3 designate each other individual into the same category.

So this set of designations satisfies all parts of ***designation symmetry***, but Rosa does not co-locate 1 & 3 (the orange cells).

So, even though there is no restriction on the eligibility of any individual for any concept, Rosa can use a preponderance methodology that satisfies ***designation unanimity***, ***designation independence***, ***designation monotonicity*** and ***non-discrimination***, but cannot use that methodology if she includes ***designation symmetry***.

9.4 The fifth principle

The fifth principle introduced by Kasher & Rubinstein gives some additional status to self-designation, and they name the principle 'the Liberal Principle'. For reasons given below, this principle is here named ***accountability***.

> ***Accountability:*** if any individual self-designates in C, Rosa does not leave C empty.

The preponderance methodology does not satisfy ***accountability*** with the profile of table 9.3 because individuals 1 and 3 self-designate in C_1, but Rosa does not locate any individual in C_1. The same is true if every individual is an $s \backslash t$ singleton, when the preponderance methodology entails that Rosa locates every individual in C_t, but every individual self-designates in C_s.

In a practical applications, ***accountability*** allows the following narrative from which its name derives[48]:

> i self-designates in C_s, but is located in C_t. i questions why this is so, and the reason is given that some other individual j has a greater claim to be located in C_s, and that i and j are sufficiently distinct that they should not be co-located.

	designation of		
designation by	i	j	others
i	C_s	C_u	any
j	C_t	C_s	
others	C_t	C_s	

*Table 9.4: **Accountability** example*

[48] Even though there is no assumption here that Rosa is in some sense accountable to the individuals in the context of the formation of her concepts.

For example, with the profile of table 9.4, i is an $s \setminus t$ singleton and no-one designates $i\&j$ in the same concept. If the weight of opinion leads Rosa to locate j in C_s, it can also support the conclusion that $i\&j$ are sufficiently different that they should not be co-located. So i would not be located in C_s despite his self-designation there. Rosa's omission of i from his self-designated concept C_s is accounted for by the stronger claim of other individuals to be located in C_s.

The name 'the Liberal Principle' for this principle seems to be less attractive, particularly because of alternative uses of the term 'liberal' in social choice, notably in the literature following Sen (1970). In that case, determining the outcome for an issue[49] is reserved for a specified individual who has the right of determination because the issue affects him personally, and any other opinion about that issue is (by implication) illiberal. This differs from the position under ***accountability***, where Rosa's decision on whether an individual is located where he self-designates is conditional on her location of other individuals.

Discussion of the applicability of ***accountability*** centres on the context. If Rosa is considering demographic concepts, or genre-musician concepts, it is arguable that she would need to justify (at least to herself) leaving a concept empty when some individual self-designates in it. Then ***accountability*** allows Rosa to justify excluding some individual from the concept in which he self-designates only when she has sufficiently strong evidence from designations that others have a prior claim to be included in that concept.

But in the context in which Rosa is categorising students into FAIL, PASS and DISTINCTION, it seems unnecessary to use a category even if an individual self-designates there. She might not think that it is reasonable to locate i in DISTINCTION where i self-designates if every other student's designation of i entails that i is in PASS or FAIL, and if all other individuals are unanimously designated in PASS or FAIL. The individual's self-designation does not have the same persuasive force in

[49] Sen's example is the question of whether a particular individual reads *Lady Chatterley's Lover* which, it is argued, should be that individual's decision alone.

determining the outcome as it might in the genre-musician or demographic contexts.

9.5 The theorem

Kasher & Rubinstein show that, when there are two concepts, the only aggregator that satisfies all of their five principles is that in which **Rosa locates each individual where he self-designates**. The principles have been modified here, but the following theorem is true for any number of categories:

> **Theorem 5.** *The only aggregator that satisfies* **designation unanimity**, **designation independence**, **designation monotonicity**, **non-discrimination** *and* **accountability** *is that in which every individual is located in the concept in which he self-designates.*

By adopting all of the five principles of theorem 5, Rosa is unable to use any other-designation to modify self-designations. Whilst this is a different conclusion from the dominant individuals entailed by theorems 2 and 4, it remains a conclusion that Rosa cannot compromise between conflicting evidence. However strong the opposition to an individual's self-designation, Rosa cannot take that opposition into account. In short, there are plausibly attractive principles that force Rosa to be liberal in her locations, and that prevent any consideration of conflicting other-designations.

9.6 Proof of theorem 5

Rosa is allocating a set I of $m \geq 3$ individuals into $2 \leq r < m$ concepts C_i. She has an aggregator that maps a profile of m^2 designations of each individual by each individual into a classification of I. The domain of the aggregator is not restricted.

The theorem is proved by showing that its negation entails a contradiction, through four stages. Between the stages some designations change.

Stage 1: Suppose that the theorem is false. Then, without losing generality, an aggregator satisfies the five principles of the theorem, and some individual h self-designates in C_1 but is located by Rosa in C_2. ***Accountability*** entails that Rosa locates some other individual(s) in C_1. These form a non-empty set $H \subset I$ and she locates each remaining individual i in some other class $C_i \neq C_1$. These designations and allocations are shown in table 9.5.

	designation of		
designation by	h	all in H	i in J
h	C_1		
members of H			
members of J			
location by Rosa	C_2	C_1	$C_i \neq C_1$
$\{h\} \cup H \cup J = I$			

Table 9.5: Proof of theorem 5 stage 1

Blank cells are unspecified, but not entirely unconstrained: ***designation unanimity*** entails that

- some individual other than h designates h in C_2;
- some individual designates $j \in H$ in C_1;
- some individual designates $i \in J$ in C_i.

Stage 2: All designations of every individual in H change (if necessary) to C_1 and all designations of each i in J change (if necessary) to C_i where C_i is defined in stage 1. These designations and locations are shown in table 9.6.

	designation of		
designation by	h	all in H	i in J
h	C_1	C_1	C_i
members of H	same as stage 1	C_1	C_i
members of J		C_1	C_i
location by Rosa	C_2	C_1	$C_i \neq C_1$

Table 9.6: Proof of theorem 5 stage 2

Designation unanimity then entails that

- Rosa continues to locate each $j \in H$ in C_1; and
- Rosa continues to locate each $i \in J$ in C_i.

No individual has changed his designation of h, and Rosa has not changed her location of any individual in $H \cup J$, so ***designation independence*** entails that Rosa continues to locate h in C_2.

Stage 3: h does not change his self-designation in C_1. Other-designations of h change (if necessary) so that every individual other than h designates h in C_2. So h is now a $1\backslash 2$-singleton. No other designation changes from stage 2. These designations and allocations are shown in table 9.7.

	designation of		
designation by	h	all in H	i in J
h	C_1	C_1	C_i
members of H	C_2	C_1	C_i
members of J	C_2	C_1	C_i
location by Rosa	C_2	C_1	$C_i \neq C_1$

Table 9.7: Proof of theorem 5 stage 3

Every individual other than h is unanimously located in stages 2 and 3, so that ***designation monotonicity*** entails that Rosa continues to locate h in C_2.

Stage 4 Designations of all individuals in H whom Rosa located in C_1 in stage 3 change so that all of those individuals become $1\backslash 2$-singletons. No other designation changes, so that h remains a $1\backslash 2$-singleton and each i in J is unanimously designated in $C_i \neq C_1$. These designations are shown in table 9.8.

		designation of		
designation by		h	members of H	all i in J
h		C_1	C_2	C_i
members of H	self	C_2	C_1	C_i
	other	C_2	C_2	C_i
members of J		C_2	C_2	C_i
location by Rosa			C_1	$C_i \neq C_1$

Table 9.8: Proof of theorem 5 stage 4

The following argument demonstrates a contradiction:

- ***Non-discrimination*** entails that Rosa co-locates all $1\backslash 2$-singletons and ***accountability*** entails that Rosa locates these individuals in C_1 because she does not locate any other individual in C_1. So Rosa locates h in C_1.
- No individual changes his designation of h between stages 3 and 4, and Rosa locates all individuals other than h in the same categories in stages 3 and 4, so ***designation independence*** entails that Rosa locates h in C_2 in stage 4 as she does in stage 3.

Starting with the assumption in table 9.5 that Rosa does not locate h where h self-designates leads to a contradiction (two conclusions on Rosa's location

of h) - so the starting assumption cannot be true. Rosa must locate every individual where he self-designates.

Note that the narrative of the proof refers to 'changes' in designations. This apparent dynamic is unnecessary (page 37). The proof rests on the assumption that every individual is eligible for every category, so that Rosa must identify locations for every individual for the sets of designations described in tables 9.5 to 9.8.

It is difficult to incorporate restrictions on eligibilities into this proof - except the trivial possibility that Rosa's set can be partitioned in a way that separates the individuals into two (or more) distinct groups: individuals in one group are eligible only for $C_1 \ldots C_q$ and those in the other group are eligible only for $C_{q+1} \ldots C_r$. Then theorem 5 can be applied within one or both groups to reach the same conclusion.

9.7 Summary

This chapter considers contexts in which the members of a community are themselves the 'objects' that are located in the concepts, and each member of the community designates himself and others to the concepts. The concepts might be demographic, rights-based as in the original application to POTENTIAL ISRAELI CITIZEN, or a temporary partition of conference delegates or students who engage in mutual evaluation.

Rosa devises her concepts using the evidence from the designations stated by each community member, and follows principles in doing so. These principles could be the same as those used in chapter 4 but these lead to the inevitability of a dominant member of the community. The alternative principles introduced here are based on those used by Kasher & Rubinstein and are different from those of chapters 4 and 7 because of the introduction here of the concept of self-designation, and its presumed difference from other-designation.

There is a simple unanimity principle, and a restricted version of a consistency principle that allows for system effects and that rules out only the possibility that Rosa relocates only a community member when only the evidence about a different member changes. These are supplemented by a monotonicity principle that prevents perverse change by Rosa when more community members support a particular designation, and by a non-discrimination principle that applies only to singletons. These are community members who self-designate in one concept about the view of all other community members who unanimously designation the singleton in another concept. The non-discrimination principle is then that Rosa colocates all similar singletons.

These four principles allow Rosa to use consistently a majority or preponderance methodology to locate each community member in line with the most common designation of that member. However, these methodologies allow that Rosa leaves some concept empty even though there are individuals who self-designate in that concept. The accountability principle prevents this - Rosa cannot locate a community member outside the concept for which that member self-designates unless she locates some other community member - who has a more persuasive 'claim' in that concept. If Rosa follows the accountability principle as well as the others, her concept formation must be based on the liberal principle that she locates every community member where that member self-designates. The context is likely to determine the reasonableness of the accountability principle.

If Rosa follows all five principles, she must follow each community member's self-designation and not use any evidence of contrary designations by others. Rose uses only one source of evidence for locating a community member in a concept just as chapters 4 and 7 conclude that, if the language is not simple, Rosa must nominate a dominant influencer, and so not use any evidence from other influencers. In both contexts, Rosa cannot compromise between those whose evidence she uses.

Chapter 10

Proof of theorems 1 and 2

10.1 Notation and conditions

10.1.1 Notation

Without specifying a context, there are:

a finite[50] set I of m individuals;

a finite set S of n objects;

a finite set Γ of r categories[51]: $\Gamma = \{C_1 \ldots C_r\}$.

$\forall C_s \in \Gamma \; E_s \subset S$ is the eligibility set for C_s, which consists of the objects that are (externally given as) eligible for C_s. Every object is in some eligibility set: $\bigcup_{i=1}^{r} E_s = S$. For a problem in which there is actual disagreement, some objects are eligible for more than one category.

$\forall s \sharp E_s \geq 2$. This facilitates the definition of linked sets (section 10.1.2).

[50] The inclusion of infinite sets in social choice generally requires the use of topological methods as in, for example, Chichilnisky & Heal (1980). That is beyond the scope here.

[51] Chapters 10 and 11 refer to categories rather than concepts, in recognition of the applicability of results in contexts other than language learning (section 1.1.1).

Profiles are $m \times n$ matrices P where "$P_{ix} = C_s$" ≡ "individual i locates x in category C_s".

The set of admissible profiles is $\Pi = \{P | \forall i \in I, x \in S\ P_{ix} = C_s \rightarrow x \in E_s\}$.

The outcome is a member of a set Q of n-vectors: $\forall q \in Q\ \forall x \in S\ q_x \in \Gamma$.

An aggregator f maps Π into Q. Each $P \in \Pi$ is mapped to a single $q \in Q$: "$q_x = f_x(P) = C_s$" ≡ "x is located in C_s in the outcome for P given aggregator f".

Where needed, $q = f(P), q' = f(P')$ etc.

For any $D \subset I$, D^c is the complement of D in I. $\sharp\Phi$ denotes the number of members of set Φ.

10.1.2 Restrictions on the domain of f

$C_s \& C_t$ are directly linked in S if $\sharp[E_s \cap E_t] \geq 2$. The assumption (section 10.1.1) that $\sharp E_s \geq 2$ entails that direct linkage is reflexive.

$C_s \& C_t$ are indirectly linked in S if there is a sequence of pairs of categories from C_s to C_t that are directly linked in S.

The domain of f is **connected** in S if all pairs of categories in Γ are directly or indirectly linked.

The domain of f is **simple** in S if $\forall x \in S\ \sharp\{s | x \in E_s\} \leq 2$.

10.1.3 Conditions on f

Class unanimity $\forall P \in \Pi\ \forall x \in S\ \forall C_s \in \Gamma\ [\forall i \in I\ P_{ix} \neq C_s \rightarrow q_x \neq C_s]$.

x is located in C_s in the outcome only if some individual locates x in C_s.

Pair unanimity (i) $\forall x, y \in S\ [\forall i \in I\ P_{ix} \neq P_{iy} \rightarrow q_x \neq q_y]$.

(ii) $\forall x, y \in S\ [\forall i \in I\ P_{ix} = P_{iy} \rightarrow q_x = q_y]$.

The outcome reflects all unanimous separate locations and all unanimous co-locations.

Class consistency $\forall P, P' \in \Pi\ \forall x \in S\ [\forall i \in I\ P_{ix} = P'_{ix} \rightarrow q_x = q'_x]$.

The outcome location of x depends only on the individuals' locations of x.

The following is straightforward.

Result 1. Class unanimity $\rightarrow \forall P \in \Pi\ \forall x \in S\ \forall C_s \in \Gamma\ [\forall i \in I\ P_{ix} = C_s \rightarrow q_x = C_s]$.

If every individual locates $x \in C_s$, then so does the outcome.

10.1.4 Responsiveness and dominance

f is **responsive to** $j \in I$ if

$\exists P, P' \in \Pi\ \left[\forall x \in S\ \forall i \neq j \in I\ [P_{ix} = P'_{ix}]\ \&\ q \neq q'\right]$.

Some change in one or more locations stated by j leads to a change in the outcome. If ***class consistency*** holds, and only j changes any locations, the outcome location of any x only changes if $P_{jx} \neq P'_{jx}$.

f is **fully responsive** if $\forall j \in I$ f is responsive to j.

j is a **dominant individual** if $\forall P \in \Pi\ \forall x \in S\ q_j = f_j(x) = P_{jx}$. Then f is responsive only to j. If ***class unanimity*** holds, $q_x = P_{jx}$ when $\forall i \in I\ P_{ix} = P_{jx}$.

Then if f is responsive only to j, and $P'_{ix} = P_{ix}$ then $q'_x = q_x = P'_{ix}$. The outcome agrees with the profile stated by j.

$\forall D \subset I, x \in S \ \forall C_s, C_t \in \Gamma$: D is **semi-dominant for** x, C_s, C_t [denoted by $sd(D; x, C_s, C_t)$] if $\forall P \in \Pi$ consistent with table 10.1, $q_x = f_x(P) = C_s$.

individuals in	
D	$P_{ix} = C_s$
D^c	$P_{ix} = C_t$

Table 10.1: Semi-dominance

$\forall D \subset I$ D is **semi-dominant** if $\forall x \in S \ \forall C_s, C_t \in \Gamma \ sd(D; x, C_s, C_t)$.

10.2 Simple domain: theorem 1

Theorem 1 states that if the domain is simple, Rosa can locate compositions in Rosa's set into genres using a fully responsive methodology if she follows deference principles ***class unanimity***, ***pair unanimity*** and ***class consistency***. More formally theorem 1 can be stated as

Formal theorem 1. *There is a fully responsive aggregator that has a simple domain and that satisfies* **class unanimity**, **pair unanimity** *and* **class consistency**.

This existence theorem can demonstrated using a simple majority aggregator and some, but not all, weighted majority aggregators.

Weighted majority aggregators are those in which a weight λ_i is assigned to each $i \in I$. Permissible weights satisfy:

$$\forall i \in I \quad 0 \leq \lambda_i < 1/2; \qquad \sum_{i \in I} \lambda_i = 1; \qquad \forall J \subset I \sum_{i \in J} \lambda_i \neq 1/2.$$

Then the weighted majority aggregator is: $\forall P \in \Pi\ \forall x \in S\ \forall C_s \in \Gamma\ \forall J \subset I\ \left[[\forall i \in J\ P_{ix} = C_s]\ \&\ \sum_{i\in J} \lambda_i > 1/2\right] \rightarrow q_x = C_s$.

The simple majority aggregator is a special case that demonstrates theorem 1 when $m = \sharp I$ is odd. Then each individual has weight $1/m$. When m is even and $j \in I$ is the tie-breaker, $\lambda_j = 2/m+1$ and $\forall i \neq j\ \lambda_i = 1/m+1$.

Result 2. *Every weighted majority aggregator satisfies* **class unanimity**, **pair unanimity** *and* **class consistency**.

Proof. $\forall x, y \in S\ (q_x = q_y = C_s) \rightarrow \sum_{i\in J} \lambda_i > 1/2$ and $\sum_{i\in J'} \lambda_i > 1/2$ where $J = \{i \in I | P_{ix} = C_s\}, J' = \{i \in I \mid P_{iy} = C_s\}$.

$$\sum_{i\in J\cup J'} \lambda_i + \sum_{i\in J\cap J'} \lambda_i = \sum_{i\in J} \lambda_i + \sum_{i\in J'} \lambda_i > 1$$

$J \cup J' \subseteq I$ so that $\sum_{i\in J\cup J'} \lambda_i \leq 1$. So $\sum_{i\in J\cap J'} \lambda_i > 0$ and $J \cap J' \neq \emptyset$. So part (ii) of ***pair unanimity*** holds. The remainder of the proof is trivial. □

The simple majority aggregator is trivially fully responsive. Not every weighted majority aggregator is fully responsive: for example, $\lambda_i = 0$ is allowed, so that the classification made by i is ignored in deriving the outcome, and then the aggregator is not responsive to i. This example extends to circumstances in which some weights are small: for example, if $\lambda_1 = \lambda_2 = 0.35, \lambda_3 = 0.2, \lambda_4 = \lambda_5 = 0.05$ the aggregator is not responsive to either individual 4 or individual 5. Adding one of these individuals to a subset of I does not increase the combined weight of members of the subset from less than $1/2$ to more than $1/2$.

10.3 Theorem 2

Theorem 2 states that if the domain is not simple and is connected, then Rosa must nominate a single dominant expert if she follows deference principles ***class***

unanimity, ***pair unanimity*** and ***class consistency*** . More formally theorem 2 can be stated as

> **Formal theorem 2.** *If an aggregator has a connected domain in S that is not simple, and satisfies* **class unanimity**, **pair unanimity** *and* **class consistency** *then there is a dominant individual.*

The proof of theorem 2 is carried through in several stages, using a route familiar from standard proofs of Arrow's theorem (chapter 12).

10.4 The epidemic

Results 3 to 9 apply $\forall D \subset I, \forall x, y \in S$, if f satisfies ***class unanimity***, ***pair unanimity*** and ***class consistency***, and if the domain is connected in S. $q = f(P), q^1 = f(P^1)$ etc.

	10.2a	10.2b		10.2c		10.2d	
individuals	P_{ix}	P_{ix}	$P_{ix'}$	P_{ix}	$P_{ix'}$	P_{ix}	P_{iy}
$i \in D$	C_s	C_s	C_s	C_t	C_s	C_s	C_u
$i \in D^c$	C_t	C_t	C_t	C_s	C_t	C_t	C_s

Table 10.2: Profile restrictions for results 3 to 6

Take an initial profile P^1 consistent with the restrictions of column 10.2a in table 10.2. Then ***class unanimity*** $\longrightarrow [q_x^1 = C_s$ or $q_x^1 = C_t]$.

Result 3. *$q_x^1 = C_s$ for the initial profile $\longrightarrow sd(D; x, C_s, C_t)$.*

Proof. Immediate from ***class consistency***. □

Result 4. *If $q_x^1 = C_s$ for the initial profile $\&\ x, x' \in E_s \cap E_t$ then $sd(D; x, C_s, C_t) \longrightarrow sd(D; x', C_s, C_t)$.*

Proof. For profiles consistent with column 10.2b: $sd(D; x, C_s, C_t) \rightarrow q_x^2 = C_s$ and ***pair unanimity*** $\rightarrow q_{x'}^2 = C_s$. Then result 3 $\rightarrow sd(D; x', C_s, C_t)$. □

Result 5. *If $q_x^1 = C_s$ for the initial profile $\& \; C_s \& C_t$ are directly linked $\& x, x' \in E_s \cap E_t, x \neq x'$ then $sd(D; x, C_s, C_t) \rightarrow sd(D; x, C_t, C_s)$.*

Proof. For profiles consistent with column 10.2c, result 4 $\rightarrow sd(D; x, C_s, C_t) \rightarrow sd(D; x', C_s, C_t) \rightarrow q_{x'}^3 = C_s$. Then ***class unanimity*** $\rightarrow q_x^3 = C_s$ or $q_x^3 = C_t$ and ***pair unanimity*** $\rightarrow q_x^3 \neq C_s$. So $q_x^3 = C_t$ and result 3 $\rightarrow sd(D; x, C_t, C_s)$. □

Result 6. *If $q_x^1 = C_s$ for the initial profile $\& x \neq y, x \in E_s \cap E_t, y \in E_s \cap E_u$ then $sd(D; x, C_s, C_t) \rightarrow sd(D; y, C_u, C_s)$.*

Proof. For profiles consistent with column 10.2d, $sd(D; x, C_s, C_t) \rightarrow q_x^4 = C_s$. Then ***class unanimity*** $\rightarrow q_y^4 = C_s$ or $q_y^4 = C_u$ and ***pair unanimity*** $\rightarrow q_y^4 \neq C_s$. So $q_y^4 = C_u$ and result 3 $\rightarrow sd(D; y, C_u, C_s)$. □

Repeated application of results 4 to 6 gives:

Result 7. (The epidemic result) $\forall D \subset I \; \forall x \in S \; C_s, C_t \in \Gamma \; [sd(D; x, C_s, C_t) \rightarrow D$ *is semi-dominant*].

To demonstrate this, there are two main scenarios to consider depending on eligibilities. Without losing generality, assume $sd(D; a, C_1, C_2)$.

10.4.1 Epidemic: scenario 1

$C_1 \& C_2$ and $C_2 \& C_3$ are directly linked. So $\exists a, a' \in E_1 \cap E_2, b, b' \in E_2 \cap E_3$ where $a \neq a', b \neq b'$. The eligibilities are illustrated in figure 10.1.

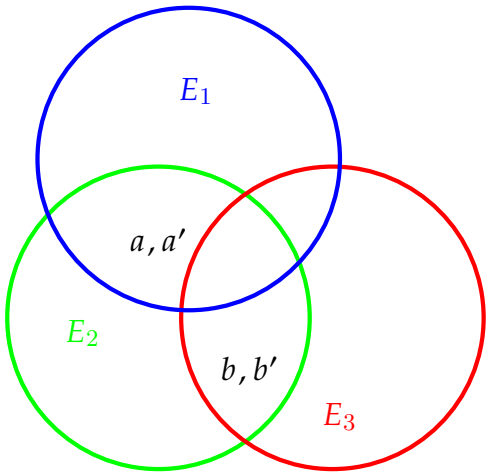

Figure 10.1: Epidemic: scenario 1

[Note that in figures of this kind, there may be other objects and other categories, subject to any eligibility restrictions]

	By assumption		$sd(D; a, C_1, C_2)$
Result 4	$s = 1, t = 2, x = a, x' = a'$	$\rightarrow$	$sd(D; a', C_1, C_2)$
Result 5	$s = 1, t = 2, x = a$	$\rightarrow$	$sd(D; a, C_2, C_1)$
Result 4	$s = 2, t = 1, x = a, x' = a'$	$\rightarrow$	$sd(D; a', C_2, C_1)$
Result 6	$s = 2, t = 1, u = 3, x = a, y = b$	$\rightarrow$	$sd(D; b, C_3, C_2)$
Result 4	$s = 3, t = 2, x = b, x' = b'$	$\rightarrow$	$sd(D; b', C_3, C_2)$
Result 5	$s = 3, t = 2, x = b'$	$\rightarrow$	$sd(D; b', C_2, C_3)$
Result 4	$s = 2, t = 3, x = b', x' = b$	$\rightarrow$	$sd(D; b, C_2, C_3)$

Table 10.3: Epidemic: scenario 1

Table 10.3 shows how the epidemic of semi-dominance spreads from $sd(D; a, C_1, C_2)$ to all other logical possibilities involving a, a', b, b' derived from figure 10.1.

10.4.2 Epidemic: scenario 2

$a \in E_1 \cap E_2$ but $C_1 \& C_2$ are not directly linked and so result 5 cannot be used directly to prove $sd(D; a, C_1, C_2) \rightarrow sd(D; a, C_2, C_1)$. Instead $\exists C_3 \in \Gamma$ such that $C_1 \& C_3$ are directly linked. Then $\exists c, c' \in E_1 \cap E_3$ where $c \neq c'$. Also $C_2 \& C_3$ are

directly or indirectly linked. The eligibilities are illustrated in figure 10.2. Because $C_1 \& C_2$ are not directly linked, result 5 cannot be used directly to demonstrate $sd(D; a, C_1, C_2) \rightarrow sd(D; a, C_2, C_1)$ Table 10.4 gives an alternative demonstration.

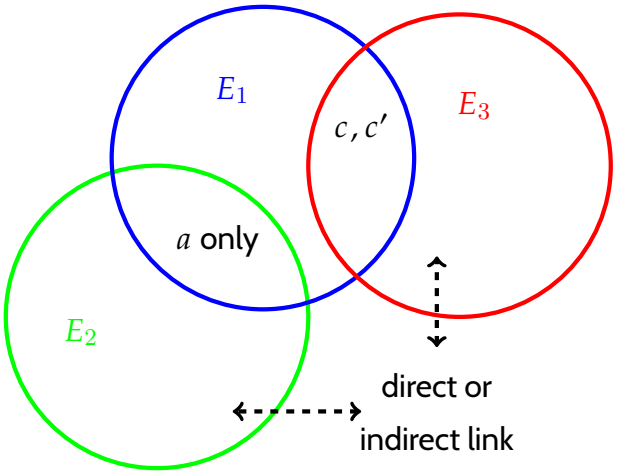

Figure 10.2: Epidemic: scenario 2

	By assumption		$sd(D; a, C_1, C_2)$
Result 6	$s = 1, t = 2, u = 3, x = a, y = c$	$\rightarrow$	$sd(D; c, C_3, C_1)$
Result 5	$s = 3, t = 1, x = c$	$\rightarrow$	$sd(D; c, C_1, C_3)$
Result 6	$s = 1, t = 3, u = 2, x = c, y = a$	$\rightarrow$	$sd(D; a, C_2, C_1)$

Table 10.4: Epidemic: scenario 2

These two examples can be combined and replicated to show the epidemic result in all circumstances.

Note that the epidemic result holds for majority aggregators in simple domains because any majority subgroup in I is semi-dominant.

The analysis of this section started with an initial profile P^1, and demonstrates that if $q_x^1 = C_s$ then D is semi-dominant. A parallel argument demonstrates that if $q_x^1 = C_t$ then D^c is semi-dominant. So

Result 8. *D is semi-dominant $\leftrightarrow$ D^c is not semi-dominant.*

Proof. By construction $q_x^1 = C_s$ or $q_x^1 = C_t$ but not both. □

Result 9. *$\forall D \subset F \subset I$ D is semi-dominant $\rightarrow$ F is semi-dominant.*

individuals	P_{ix}^5	P_{iy}^5
in D	C_2	C_3
in F^c	C_1	C_2
all others	C_1	C_3

Table 10.5: profile restrictions of result 9

Proof. Without losing generality assume that $C_1 \& C_2$ and $C_2 \& C_3$ are directly connected so that $\exists x \in E_1 \cap E_2$ $y \in E_2 \cap E_3$. Suppose that D is semi-dominant but $F \supset D$ is not semi-dominant. Then by result 8 F^c is semi-dominant. So for a profile P^5 consistent with table 10.5 $q_x^5 = q_y^5 = C_2$ contrary to ***pair unanimity***. So F is semi-dominant. □

10.5 Intersection

Result 10. *Under the conditions of theorem 2, if f is an aggregator such that $D_1 \& D_2$ are semi-dominant then $D \cap D' \neq \emptyset$.*

Proof. If D_1 is semi-dominant, result 8 entails D_1^c is not semi-dominant. But $D_1 \cap D_2 = \emptyset \rightarrow D_2 \subset D_1^c$ so if D_2 is semi-dominant, result 9 $\rightarrow$ D_1^c is semi-dominant which is a contradiction □

Then

Result 11. *Under the conditions of theorem 2, if $D_1 \neq D_2 \subset I$ are both semi-dominant then $D_1 \cap D_2$ is semi-dominant.*

There are several possible scenarios, illustrated in figures 10.3 to 10.6. These combine a variety of direct and indirect links and can be extended and adapted to other scenarios, including those involving additional objects and classes from those that are explicitly mentioned. The important distinction between scenario 1 and others is that, in scenario 1, there are two objects that are eligible for three categories; the definition of a non-simple domain applies when only one object is eligible for three categories. This possibility is shown in scenarios 2 to 4.

10.5.1 Intersection: scenario 1

$v \neq w \in E_1 \cap E_2 \cap E_3$ so that $C_1 \& C_2$, $C_1 \& C_3$ and $C_2 \& C_3$ are directly linked. This is illustrated in figure 10.3 and profile restrictions consistent with the eligibilities are in table 10.6.

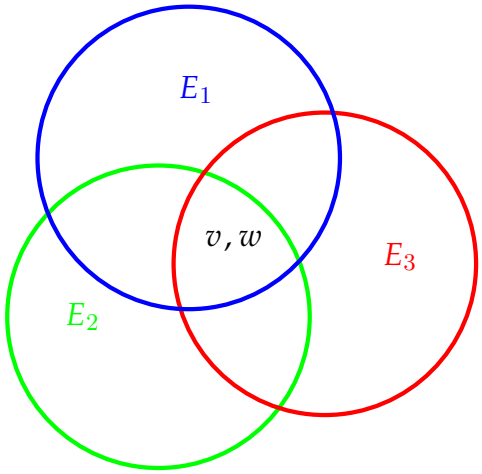

Figure 10.3: Intersection: scenario 1

individuals	10.6a		10.6b		10.6c	
in	P_{iv}	P_{iw}	P'_{iv}	P'_{iw}	P''_{iv}	P''_{iw}
$D_1 \cap D_2$	C_1	C_2	C_1	C_3	C_1	C_2
$D_1 \cap D_2^c$	C_3	C_2	C_3	C_2	C_3	C_1
$D_1^c \cap D_2$	C_2	C_1	C_2	C_3	C_2	C_1
$D_1^c \cap D_2^c$	C_3	C_1	C_3	C_2	C_3	C_1

Table 10.6: Profile restrictions intersection, scenario 1

- For profiles consistent with column 10.6a: D_1 is semi-dominant $\rightarrow q_w = C_2$. ***Pair unanimity*** $\rightarrow q_v \neq C_2$.
- For profiles consistent with column 10.6b: D_2 is semi-dominant $\rightarrow q'_w = C_3$. ***Pair unanimity*** $\rightarrow q'_v \neq C_3$.
- ***Class consistency*** $\rightarrow q_v = q'_v = q''_v$, and so ***class unanimity*** $\rightarrow q''_v = C_1$.
- ***Pair unanimity*** $\rightarrow q''_w \neq C_1$. ***Class unanimity*** $\rightarrow q''_w = C_2$ and so
- Result 2 $\rightarrow sd(D_1 \cap D_2; w, C_2, C_1)$ and result 7 $\rightarrow D_1 \cap D_2$ is semi-dominant.

10.5.2 Intersection: scenario 2

In this scenario, $E_1 \& E_2$ and $E_2 \& E_3$ are directly linked.

$E_1 \cap E_2 \cap E_3 = \{w\}, v \in E_1 \cap E_2, z \in E_2 \cap E_3$ so that $C_1 \& C_2$, $C_2 \& C_3$ and $C_1 \& C_3$ are indirectly linked. This is illustrated in figure 10.4.

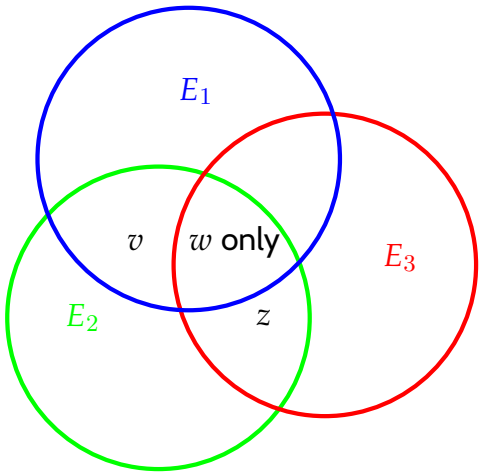

Figure 10.4: Intersection: scenario 2

10.5.3 Intersection: scenario 3

In this scenario, $E_1 \& E_2$ are directly linked and $E_2 \& E_3$ are indirectly linked. The scenario can be adapted if $E_1 \& E_2$ are indirectly linked and $E_2 \& E_3$ are directly linked.

$E_1 \cap E_2 \cap E_3 = \{w\}, v \in E_1 \cap E_2, z, z' \in E_3 \cap E_4, C_2 \& C_4$ are directly or indirectly linked. So all pairs of categories in $C_1 \ldots C_4$ are directly or indirectly linked. This is illustrated in figure 10.5.

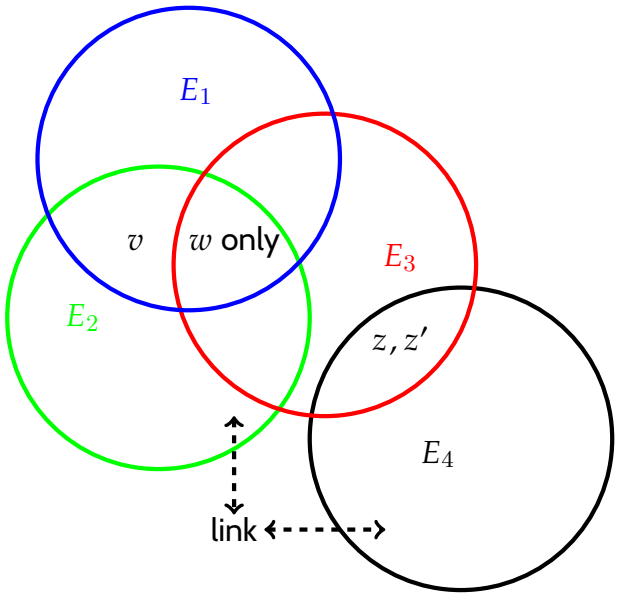

Figure 10.5: Intersection: scenario 3

10.5.4 Intersection: scenario 4

In this scenario both $E_1 \& E_2$ and $E_2 \& E_3$ are indirectly linked.

$E_1 \cap E_2 \cap E_3 = \{w\}, v, v' \in E_2 \cap E_5, z, z' \in E_3 \cap E_4, C_1 \& C_5$ and $C_2 \& C_4$ are directly or indirectly linked. So all pairs of categories in $C_1 \ldots C_5$ are directly or indirectly linked. This is illustrated in figure 10.6.

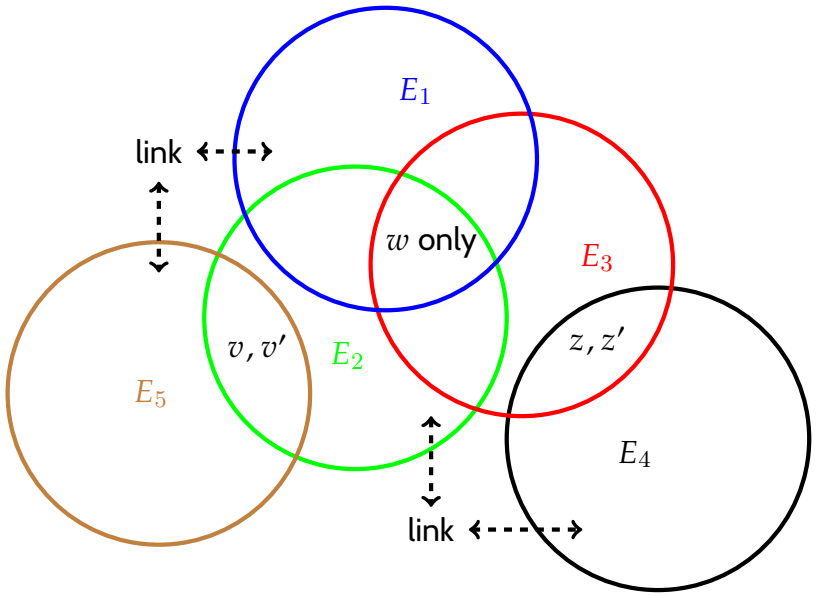

Figure 10.6: Intersection: scenario 4

individuals	10.7a			10.7b		
in	P_{iw}	P_{iv}	P_{iz}	P'_{iw}	P'_{iv}	
$D_1 \cap D_2$	C_1	C_2	C_3	C_1	C_2	For scenario 2, set $C_s = C_1, C_t = C_2$
$D_1 \cap D_2^c$	C_3	C_2	C_t	C_3	C_s	For scenario 3, set $C_s = C_1, C_t = C_4$
$D_1^c \cap D_2$	C_2	C_s	C_3	C_2	C_s	For scenario 4, set $C_s = C_5, C_t = C_4$
$D_1^c \cap D_2^c$	C_3	C_s	C_t	C_3	C_s	

Table 10.7: Profile restrictions intersection, scenarios 2,3,4

- For profiles consistent with column 10.7a: D_1 is semi-dominant $\longrightarrow q_v = C_2$. ***Pair unanimity*** $\longrightarrow q_w \neq C_2$.
- For profiles consistent with column 10.7a: D_2 is semi-dominant $\longrightarrow q_z = C_3$. ***Pair unanimity*** $\longrightarrow q_w \neq C_3$.
- ***Class consistency*** $\longrightarrow q_w = q'_w$, and so ***class unanimity*** $\longrightarrow q''_w = C_1$.
- ***Pair unanimity*** $\longrightarrow q''_v \neq C_1$. ***Class unanimity*** $\longrightarrow q''_v = C_2$ and so
- Result 2 $\longrightarrow sd(D_1 \cap D_2; v, C_2, C_1)$ and result 7 $\longrightarrow D_1 \cap D_2$ is semi-dominant.

Taken together the four scenarios are exhaustive of the possibilities, giving a proof of result 11.

Note that result 11 does not hold when the domain is simple. For example, with the majority aggregator, any majority set is semi-dominant. But the intersection of two majority sets is not necessarily a majority set, and so is not semi-dominant.

10.6 The dominant individual

Define D^* as the semi-dominant subset of I with the fewest members. The following results hold given conditions of theorem 2.

Result 12. D^* *is unique*

Proof. If $D^{**} \neq D^*$ and $\sharp D^{**} = \sharp D^*$, then result 11 $\rightarrow D^{**} \cap D^*$ is semi-dominant and $\sharp(D^{**} \cap D^*) < \sharp D^*$ contradicting the construction of D^*. □

Result 13. $\sharp D^* = 1$

Proof. Suppose that $j \neq k \in D^*$. Partition I into $F \& F^c$ where $j \in F, k \in F^c$. Then, $k \notin F \rightarrow F \cap D^*$ is a proper subset of D^*, so $F \cap D^*$ is not semi-dominant and result 11 $\rightarrow F$ is not semi-dominant. So result 10 $\rightarrow F^c$ is semi-dominant, and result 11 $\rightarrow F^c \cap D^*$ is semi-dominant. But $\sharp(F^c \cap D^*) < \sharp D^*$ because $j \notin F^c$, contradicting the construction of D^*. So $\sharp D^* = 1$. □

So, under the conditions of theorem 2, the smallest semi-dominant set has a single member.

Result 14. *Under the conditions of theorem 2,* $D^* = \{k\} \rightarrow k$ *is a dominant individual.*

	10.8a	10.8b	
individuals	P^1_{ix}	P^2_{ix}	P^2_{iy}
k	C_1	C_1	C_2
D_1	C_1	C_1	C_t
D_2	C_2	C_2	C_t
⋮	⋮	⋮	⋮
D_t	C_t	C_t	C_2
⋮	⋮	⋮	⋮
D_r	C_r	C_r	C_t
$\{k\} \cup \bigcup_{s=1}^{r} D_s = I$			
$\forall s \; x \notin E_s \rightarrow D_s = \emptyset$			

Table 10.8: D^* consists of a dominant individual

Proof. If the result is false, $\exists, C_1, C_2 \in \Gamma \; x \in S \; P^1 \in \Pi$ consistent with 10.8a such that $q^1 = C_2$. For some[52] $C_t \in \Gamma \; C_2 \& C_t$ are directly connected so $\exists y \in S \; y \in E_2 \cup E_t$. For $P^2 \in \Gamma$ consistent with 10.8b:

- ***class independence*** $\rightarrow q^2_x = C_2$;
- Result 9 & $\{k\}$ is semi-decisive $\rightarrow \{k\} \cup D_t$ is semi-decisive $\rightarrow q^2_y = C_2$.

which contradicts ***pair unanimity***. So the result is true. □

10.7 Domain not connected

Theorem 1 does not require that the domain is connected because a majority methodology can define an aggregator that can be used for all objects, whether or not categories are linked. Theorem 2 relies on a connected domain to prove the epidemic proposition and to limit the configurations to those of figures 10.3 to 10.6.

[52] It is possible that $D_t = \emptyset$ which is certainly the case when $y \notin E_1$. If $y \in E_1$ then t can be set equal to 1.

In principle there are domains that are not connected and not simple. This section considers two of these.

10.7.1 A separable domain

In the musical genres example, the domain might be separable into two or more wholly unconnected parts. CLASSICAL, OPERA and JAZZ are typical genres of BBC Radio 3, while Radio 1 specialises in POP and related genres such as HIP-HOP and RAP. Arguably there is no overlap (in terms of eligibilities) between these two sets of genres.

In formal terms,

T, U form a partition of S, and we say that $T\&U$ are **separable** just if $\forall C_s, x \in T, y \in U\ [x \in E_s \Leftrightarrow y \notin E_s]$.

A category for which objects in T are eligible and a category for which objects in U are eligible are not linked, directly or indirectly. Then it is trivial that a different methodology can be used for determining the outcome locations of objects in T from that used for objects in U. Depending on the configuration (and assuming that pairs of categories within the two subsets of Γ are directly or indirectly linked) possible aggregators are then

- objects in the two subsets can be located using possibly different weighted majority methodologies;
- objects in one subset are located by a dominant individual; objects in the other can be located using a weighted majority methodology;
- objects in the two subsets are located by possibly different dominant individuals.

This conclusion can be extended if S can be partitioned into more than two separable subsets.

10.7.2 Other non-connected domains

There are possible configurations in which the domain is not connected in S, but S cannot be partitioned into two or more separable subsets. Some of these might lead to the same conclusion as theorem 2, but the following example shows that this is not necessarily the case.

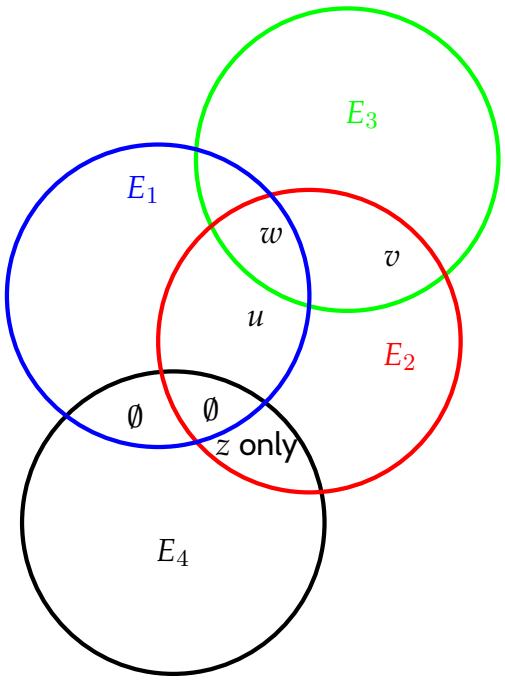

Figure 10.7: Non-linked classes

In figure 10.7, the domain is not simple because $w \in E_1 \cap E_2 \cap E_3$. $C_1 \& C_2$ and $C_2 \& C_3$ are directly linked, but C_4 is not directly linked to any other category. S cannot be partitioned into separable subsets because $u, v, w, z \in E_2$.

Theorem 2 entails that some individual is dominant for the location of $u, v \& w$. But it is possible to define an aggregator that satisfies ***class unanimity***, ***pair unanimity*** and ***consistency*** as follows: $\forall P \in \Pi$

$q_u = P_{1u}, q_v = P_{1v}, q_w = P_{1w}$;

$\forall i \in I \; P_{iz} = C_2 \rightarrow q_z = C_2$; otherwise $q_z = C_4$.

where $q = f(P)$. Individual 1 determines the location of objects other than z, but $z \in C_4$ unless every individual locates $z \in C_2$.

individual	P_{iz}	P_{iu}	P_{iv}	P_{iw}
1 (dominant for u, v, w)	C_2			
some others	C_4			
remainder	C_2			
outcome	C_4	P_{1u}	P_{1v}	P_{1w}

Table 10.9: Non-linked classes

Then, with a profile that is consistent with the restrictions of table 10.9, $q_z = C_4 \neq P_{1z}$. No individual is dominant. Theorem 2 does not hold if the domain of the aggregator is not connected.

The fact the domain is not simple entails that there can be profiles for which a majority aggregator cannot be used. In this example there is an alternative aggregator that does not involve a dominant individual that satisfies all the conditions of theorem 2 except that the domain is not connected. Arguably in the musical genres application, this domain is coincidental because it so happens that S (Rosa's set) only contains one composition that is common to E_4 and any other eligibility set. If S were to expand to include other compositions, that uniqueness might disappear, so that the only available aggregator that satisfies all the conditions would involve a dominant individual. The rigorous application of the techniques of social choice theory must allow for such idiosyncrasies, but arguably they are unimportant in our main context.

10.8 Summary

This chapter follows a traditional route in social choice theory to prove the results introduced in chapter4. That route involves the demonstration that, under the specified conditions, if the outcome follows a subset of individuals to determine

one location, it must follow that subset whenever they express a unanimous view. This is the epidemic. The traditional route then considers the intersection of these semi-dominant sets, and finally shows that the smallest such semi-dominant set contains only one dominant individual.

The chapter concludes with a formal demonstration that theorem 2 does not hold if the domain is not connected. There is then an alternative aggregator that satisfies all the principles and that does not involve a dominant individual.

Chapter 11

Proof of theorems 3 and 4

11.1 Introduction

This chapter includes confirmation of the use of a majority-based aggregator in a simple graded domain (theorem 3, page 110), and a proof of theorem 4. The proof strategy for the latter is the same as that for theorem 2 in sections 10.4 to 10.6, including an epidemic and an intersection result. However, the conditions on the aggregator and the proof emphasise the placement of boundaries between categories by individuals and the outcome.

11.1.1 Notation

As in chapter 10, there are a finite set I of m individuals, a finite set S of n objects and a finite set Γ of r categories: $\Gamma = \{C_1 \ldots C_r\}$. Each C_s has an eligibility set E_s. In this chapter, the objects in S are graded by a strict ordering $\succ$, and for convenience of notation, $y \succeq x \Leftrightarrow [y \succ x$ or $y \equiv x]$. $y \succ x$ are neighbouring objects in S, written as $N(x, y)$ just if $\nexists z \in S \; y \succ z \succ x$. We write $N(w, x, y) \iff N(w, x)$ & $N(x, y)$. We use the terminology that if $y \succ x$ then y is above x and x is below y.

$\bigcup_{i=1}^{r} E_s = S$, so that every object is in at least one eligibility set. Every eligibility set is convex using $\succ$ so that $E_s = \{a_s \ldots b_s\}$. $\forall s\ a_s \succ a_{s-1}\ \&\ b_s \succ b_{s-1}\ \&\ b_s \succ a_{s+1}$. The last condition entails $b_s \not\equiv a_{s+1}$ and $\sharp(E_s \cap E_{s+1}) \geq 2$. This is the equivalent in the graded context of the assumption of a connected domain in section 10.1.2.

11.1.2 Boundary placement

The challenge of locating objects in categories in a graded context is equivalent to the challenge of placing boundaries between objects. There are $n-1$ distinct boundary placements $\langle x, y\rangle$ where $N(x, y)$, and each individual $i \in I$ places $r-1$ boundaries β_{is} for $s = 1 \ldots r-1$. β_{is} is i's placement of the boundary between $C_s \& C_{s+1}$. With an obvious extension of the use of $\succ$, $\forall i, s\ \beta_{i,s+1} \succ \beta_{is}$. If $N(x, y)\ \&\ y \succ \beta_{is} \succ x$, we write $y \blacktriangleright \beta_{is} \blacktriangleright x$. $\blacktriangleright$ then denotes the relation "is just above".

The **archetype assumption** (page 103) is that $\exists x \in S\ \sharp\{s : x \in E_s\} = 1 \leftrightarrow E_{s-1} \cap E_s \cap E_{s+1} = \emptyset$. This is illustrated by the configurations of figures 7.4 and 7.5.

F_s is the set of eligible placements for β_{is}. $\forall s\ F_s = \{\langle x, y\rangle : x \in E_s, y \in E_{s+1}, N(x, y)\}$.

A profile P is an $m \times (r-1)$ matrix of boundary placements, one for each individual. The set of admissible profiles Π contains all and only profiles in which $\forall i, s\ \beta_{is} \in F_s$ and in which **no category is empty in any individual classification**: $\forall i, s\ \exists x \in S\ \beta_{i,s+1} \succ x \succ \beta_{is}$. This entails that no two boundaries are placed together.

An aggregator f has domain Π and for any $P \in \Pi$ defines an (r-1)-vector $f(P)$ of **outcome boundary placements** β_{fs}.

11.1.3 Simple and non-simple domains

The domain of f is **simple** in a graded context if $\forall s\ F_s \cap F_{s+1} = \emptyset$. Then there is no possible place for which more than one boundary is eligible. By extension, the domain of f is **not simple** if $\exists s\ F_s \cap F_{s+1} \neq \emptyset$.

Figures 11.1 and 11.2 are used for illustration here.

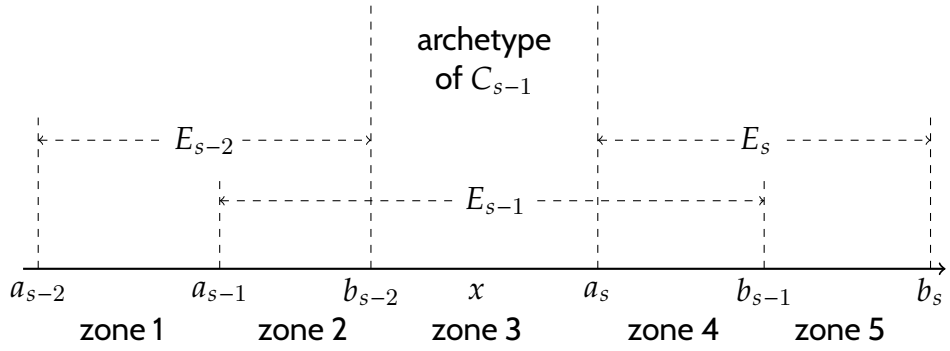

Figure 11.1: Illustrative configuration A

In figure 11.1 the domain can be simple only because we make the archetype assumption that $\exists x\ a_s \succ x \succ b_{s-2}$ in zone 3.

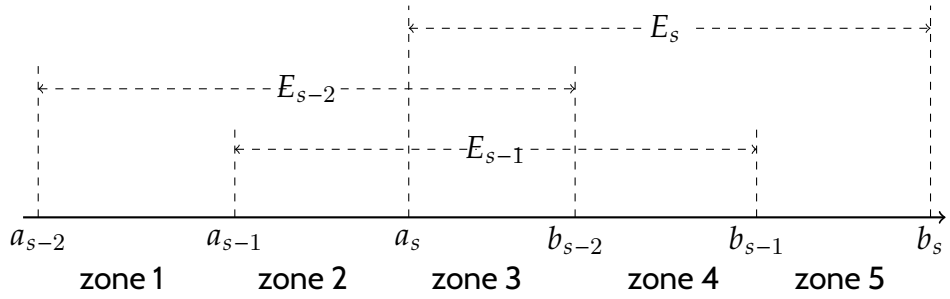

Figure 11.2: Illustrative configuration B

Figure 11.2 can be modified to allow $a_s \equiv b_{s-2}$. Then $\sharp(E_{s-2} \cap E_{s-1} \cap E_s) = 1$.

11.2 Conditions on f

In these definitions, $P = [\beta_{is}], P' = [\beta'_{is}], f(P) = (\beta_{fs}), f(P') = (\beta'_{fs})$

Boundary unanimity (i) $\forall P \in \Pi\ \forall x, y\ \forall s\ \big[[\forall i\ \neg(y \succ \beta_{is} \succ x)] \rightarrow \neg(y \succ \beta_{fs} \succ x)\big]$.

(ii) $\forall P \in \Pi\ \forall y\ \forall s\ max_i \sharp\{s : y \succ \beta_{is}\} \geq \sharp\{s : y \succ \beta_{fs}\} \geq min_i \sharp\{s : y \succ \beta_{is}\}$

Boundary consistency (i) $\forall P, P' \in \Pi\ \forall x\ \forall s \big[\forall i \in I\ [\beta_{is} \succ x \leftrightarrow \beta'_{is} \succ x] \rightarrow [\beta_{fs} \succ x \leftrightarrow \beta'_{fs} \succ x]\big]$.

(ii) $\forall P, P' \in \Pi\ \forall x\ \big[\forall i \in I\ [\exists s\ \beta_{is} \blacktriangleright x \leftrightarrow \exists t\ \beta'_{it} \blacktriangleright x] \rightarrow [\exists s\ \beta_{fs} \blacktriangleright x \leftrightarrow \exists t\ \beta'_{ft} \blacktriangleright x]\big]$.

Boundary monotonicity $\forall J, K \subset I\ \forall s, x\ \forall P, P' \in \Pi$ satisfying the restrictions of table 11.1, then $[\beta_{fs} \blacktriangleright x] \rightarrow [\beta'_{fs} \blacktriangleright x]$.

$i \in$	P	P'
J	$\beta_{is} \blacktriangleright x$	$\beta'_{is} \blacktriangleright x$
K	$x \blacktriangleright \beta_{is}$	$\beta'_{is} \blacktriangleright x$
$(J \cup K)^c$	$x \blacktriangleright \beta_{is}$	$x \blacktriangleright \beta'_{is}$

Table 11.1: Boundary monotonicity

Boundary symmetry $\forall J \subset I\ \forall s, x\ \forall P, P' \in \Pi$ satisfying the restrictions of table 11.2, then $[\beta_{fs} \blacktriangleright x] \leftrightarrow [x \blacktriangleright \beta'_{fs}]$.

$i \in$	P	P'
J	$\beta_{is} \blacktriangleright x$	$x \blacktriangleright \beta'_{is}$
J^c	$x \blacktriangleright \beta_{is}$	$\beta'_{is} \blacktriangleright x$

Table 11.2: Boundary symmetry

11.3 Theorems 3 and 4

As in chapter 10, the domain of f is simple if $\forall s, F_s \cap F_{s+1} = \emptyset$. At most one boundary is eligible to be placed in each of the $n-1$ possible boundary placements. Otherwise the domain is not simple. Also, we assume that $\forall s \ \sharp F_s > 1$ so that the individuals can disagree about the location of each boundary. Then we have

Formal theorem 3. *There is a fully responsive aggregator that has a simple domain and that satisfies* **boundary unanimity**, **boundary consistency**, **boundary monotonicity** *and* **boundary symmetry**.

The proof of this theorem follows in a similar way to the proof of theorem 1, using a majority aggregator.

Formal theorem 4. *If an aggregator f has a domain that is not simple, and satisfies* **boundary unanimity**, **boundary consistency**, **boundary monotonicity** *and* **boundary symmetry** *then f entails that there is a dominant individual.*

11.4 Proof of theorem 4

The results of this section and of section 11.4.2 are proved under the conditions used in theorem 4. We also use the labelling a_s, b_s etc shown in figures 11.1 and 11.2.

11.4.1 Semi-dominance and the epidemic

In section 10.1.4, $D \subset I$ is semi-dominant for x, C_s, C_t if the outcome locates $x \in C_s$ whenever all members of D locate $x \in C_s$ and all members of D^c locate $x \in C_t$. In a graded context, we need to consider semi-dominance only when $t = s + 1$. Then D is semi-dominant for $x\&C_s$ [written as $sd(D; x, s)$] if $\beta_{fs} \blacktriangleright x$ for all $P \in \Pi$ consistent with column 11.3a in table 11.3.

$i \in$	11.3a	11.3b
D	$\beta_{is} \blacktriangleright x$	$\beta_{is} \succ x$
D^c	$x \blacktriangleright \beta_{is}$	$x \succ \beta_{is}$

Table 11.3: Profile restrictions 11.3a& b

D is semi-dominant if $\forall x, s\ sd(D; x, s)$

Take some initial profile that is consistent with column 11.3a in table 11.3. ***Boundary unanimity (i)*** entails that $\beta_{fs} \blacktriangleright x$ or $x \blacktriangleright \beta_{fs}$ for any profile consistent with 11.3a. If we make the **initial assumption** that $\beta_{fs} \blacktriangleright x$ for the initial profile then ***boundary consistency (i)*** entails $sd(D; x, s)$. Otherwise $sd(D^c; w, s)$ where $N(w, x)$. Furthermore, ***boundary consistency (i)*** and $sd(D; x, s)$ together entail $\beta_{fs} \succ x$ for all profiles consistent with column 11.3b.

Result 15. $\forall y \succeq a_{s+1} \in F_s, x \in F_s\ sd(D; x, s) \rightarrow sd(D; y, s)$.

$i \in$	11.4a	11.4b
D	$\beta_{is} \blacktriangleright y \succ x$	$\beta_{is} \blacktriangleright y \succ x$
D^c	$y \succ x \blacktriangleright \beta_{is}$	$y \blacktriangleright \beta_{is} \succ x$

Table 11.4: Profile restrictions for result 15

Proof. For any profile consistent with column 11.4a $sd(D; x, s)$ & ***boundary consistency (i)*** & ***boundary unanimity (i)*** together entail $\left[\beta_{fs} \succ x\ \&\neg[y \succ \beta_{fs} \succ x]\right]$ so that $\beta_{fs} \blacktriangleright y$. Then for any profile consistent with 11.4b, ***boundary consistency (i)*** $\rightarrow \beta_{fs} \blacktriangleright y \rightarrow sd(D; y, s)$. A similar argument shows $sd(D; w, s)$ for all $x \succ w \succeq a_{s+1}$ □

The following is trivial.

Result 16. Boundary symmetry *&* $sd(D, x, s)$ *together entail* $x \blacktriangleright \beta_{fs}$ *for any profile consistent with the restrictions of table 11.5.*

$i \in$	restriction
D	$x \blacktriangleright \beta_{is}$
D^c	$\beta_{is} \blacktriangleright x$

Table 11.5: Profile restriction for result 16

There are then two possibilities, illustrated in figures 11.1 and 11.2.

Result 17. *In configuration A (figure 11.1),* $sd(D; b_s, s) \rightarrow sd(D; a_{s-2}, s-1)$

$i \in$	restriction
D	$\beta_{is} \blacktriangleright a_s \succ \beta_{i,s-1} \blacktriangleright b_{s-2}$
D^c	$a_s \blacktriangleright \beta_{is} \succ b_{s-2} \blacktriangleright \beta_{i,s-1}$

Table 11.6: Profile restrictions for result 17

Proof. If the result is false, then for some profile consistent with the restriction of table 11.6, $b_{s-2} \blacktriangleright \beta_{f,s-1}$. But $sd(D; a_s, s) \rightarrow \beta_{fs} \blacktriangleright a_s \rightarrow \beta_{fs} \blacktriangleright a_s \succ b_{s-2} \blacktriangleright \beta_{f,s-1}$. Each individual places one boundary between $b_{s-2}\&a_s$ but there is no boundary in the outcome between $b_{s-2}\&a_s$ violating ***boundary unanimity (ii)***. So the result is true. □

Result 18. *In configuration B (figure 11.1),* $sd(D; b_s, s) \rightarrow sd(D; b_s, s-1)$.

$i \in$	profile P	profile P'
D	$\beta_{is} \blacktriangleright b_s \succ \beta_{i,s-1}$	$\beta'_{is} \succ \beta'_{i,s-1} \blacktriangleright b_s$
D^c	$b_s \blacktriangleright \beta_{is} \succ \beta_{i,s-1}$	$\beta'_{is} \succ b_s \blacktriangleright \beta'_{i,s-1}$

Table 11.7: Profile restrictions for result 18

Proof. For $P, P' \in \Pi$ consistent with table 11.7 $\forall i \in I$ $\beta_{is} \blacktriangleright b_s \leftrightarrow \beta'_{i,s-1} \blacktriangleright b_s$. $sd(D; b_s, s) \rightarrow \beta_{fs} \blacktriangleright b_s$. Then ***boundary consistency (ii)*** $\rightarrow \beta'_{f,s-1} \blacktriangleright b_s$ and hence $sd(D; b_s, s-1)$ □

Results 17 and 18 show how the epidemic spreads from the placement of β_{fs} to the placement of $\beta_{f,s-1}$. A parallel argument can be used to show how the epidemic spreads from the placement of β_{fs} to the placement of $\beta_{f,s+1}$. Then sequential use of results 15, 17 and 18 demonstrate the full epidemic of semi-dominance:

Result 19. $\exists x, s\ sd(D; x, s) \rightarrow D$ *is semi-dominant.*

If the initial assumption is false, D^c is semi-dominant, so that, in parallel with results in section **??** (page **??**):

Result 20. D *is semi-dominant* $\leftrightarrow D^c$ *is not semi-dominant.*

and from ***boundary monotonicity***

Result 21. *If* D *is semi-dominant and* $D' \supset D$ *then* D' *is semi-dominant.*

and

Result 22. *If* $D_1 \& D_2$ *are semi-dominant then* $D_1 \cap D_2 \neq \emptyset$.

11.4.2 Intersection

Result 23. *If* $D_1 \& D_2$ *are semi-dominant then* $D_1 \cap D_2$ *is semi-dominant.*

Proof. By result 21, if D_2 is semi-dominant then $D_2 \cup (D_1^c \cap D_2^c) = (D_2 \cap D_1) \cup (D_2 \cap D_1^c) \cup (D_1^c \cap D_2^c) = (D_2 \cap D_1) \cup D_1^c$ is semi-dominant.

The domain of f is not simple, and, without losing generality, suppose that $N(w, x, y), \langle w, x \rangle \in F_{s-1} \cap F_s$.

$i \in$	profile P	profile P'
$D_1 \cap D_2$	$y \blacktriangleright \beta_{is} \blacktriangleright x \blacktriangleright \beta_{i,s-1} \blacktriangleright w$	$y \blacktriangleright \beta'_{is} \blacktriangleright x \blacktriangleright \beta'_{i,s-1} \blacktriangleright w$
$D_1 \cap D_2^c$	$\beta_{is} \blacktriangleright y \succ x \blacktriangleright \beta_{i,s-1} \blacktriangleright w$	$y \succ x \blacktriangleright \beta'_{is} \blacktriangleright w \blacktriangleright \beta'_{i,s-1}$
D_1^c	$y \blacktriangleright \beta_{is} \blacktriangleright x \succ w \blacktriangleright \beta_{i,s-1}$	$y \blacktriangleright \beta'_{is} \blacktriangleright x \succ w \blacktriangleright \beta'_{i,s-1}$

Table 11.8: Profile restrictions for result 23

Then

- D_1 is semi-dominant $\rightarrow x \blacktriangleright \beta_{f,s-1} \blacktriangleright w$
- $(D_2 \cap D_1) \cup D_1^c \supseteq D_2$ is semi-dominant $\rightarrow y \blacktriangleright \beta'_{fs} \blacktriangleright x$
- ***boundary consistency (ii)*** $\rightarrow x \blacktriangleright \beta'_{f,s-1} \blacktriangleright w$ or $x \blacktriangleright \beta'_{fs} \blacktriangleright w$.
- $y \blacktriangleright \beta'_{fs} \blacktriangleright x \rightarrow \neg(x \blacktriangleright \beta'_{fs} \blacktriangleright w)$
- So $x \blacktriangleright \beta'_{f,s-1} \blacktriangleright w$ and ***boundary consistency (i)*** $\rightarrow sd(D_1 \cap D_2; w, s-1)$
- Then the epidemic result 19 entails that $D_1 \cap D_2$ is semi-dominant.

□

The remainder of the proof of theorem 4 follows as it does for theorem 2 in result 14 on page 165 and in section 10.6.

11.5 Summary

This chapter follows a similar strategy to that of chapter 10 to prove the results stated in chapter 7 for a graded context. Although all the conditions on the placement of boundaries are used in the proof, no attempt is made to demonstrate that there are alternative, non-dominant, aggregators available if one condition is removed.

Chapter 12

Social choice

12.1 Introduction

Theorems 2, 4 and 5 all have similarities to those that are familiar in social choice. They show that, in different contexts, it is not possible to find a compromise between conflicting views if the range of disagreement is sufficiently large, and if reasonable sounding conditions (principles, axioms …) are imposed. Theorems 1 and 3 show domain restrictions that avoid the dictatorial outcomes of theorems 2 and 4 by permitting the consistent use of a majority-based aggregator.

The potential for a dictatorial outcome arises in many different social choice **structures** - where a structure includes the domain of evidence to be aggregated, the form of the outcome of the aggregation, and the conditions imposed on the aggregator. There is a significant literature on domain restrictions that permit the use of a majority-based aggregator in the Arrow structure which concerns the aggregation of orderings of a set of alternatives. Sections 12.2 to 12.4 outline this and two other structures. Section 12.5 compares these structures with those that are introduced in chapters 4 and 7.

These are all particular applications of the general structures of Wilson (1975) and Rubinstein & Fishburn (1986) in which the conditions imposed on the aggregator are (or include) unanimity and independence, and the conclusion is that, with an unrestricted domain, the only available aggregators are dictatorial.

12.2 The Arrow structure

The original and most widely developed structure in social choice is that of Arrow (1951). This involves the aggregation of orderings of objects to produce an outcome that is also an ordering (or sometimes only a 'winner'). Its principal conditions can be stated in several forms, of which a common version includes unanimity (if every individual places x above y, then so does the outcome) and an independence of irrelevant alternatives condition (the outcome order of x and y depends only on the individual orders of x and y). The dictatorial outcome arises if the domain of the aggregator is unrestricted, allowing any individual to order the objects in any logically possible way. This structure is **binary** in that the evidence and outcome are binary relations (orderings) and the conditions are expressed in terms of components of the binary relation. Many proofs of Arrow's theorem follow the same route as the proof of theorem 2 in chapter 10. There is an epidemic proposition involving decisive (or dominant) sets of individuals, a result on the intersection of decisive sets, and a demonstration that the smallest decisive set has a single member.

Arrow's theorem is often illustrated in the context of elections and committee decisions. This context can be traced back to Condorcet (1785), Nanson (1882), Black (1948) and many others. Many of the influences on Arrow's work came from traditions in welfare economics that try to identify optimal economic policies, but that do not have a rigorous basis such as that developed in his 1951 book. Arrow and others developed (for example, Arrow & Hahn, 1971, Debreu, 1959) a rigorous basis for the analysis of market interrelations (general equilibrium theory), and the rigour of that was translated to social choice. The axiomatic 'model-building' approach

that Arrow pioneered is a fundamental component of social choice theory as it has developed since 1951.

Arrow's structure has applications in ethics. At an individual level, the long-standing utilitarian tradition of measuring and adding individual utilities ('the greatest good of the greatest number') to establish the best alternative of several under consideration involves difficulties both of measurement and of interpersonal comparison. Social choice theory aims to avoid these difficulties by recognising that an individual's judgement will be based on her aggregation of the different impacts of the alternatives on the people whose views that she wants to incorporate.

12.2.1 Domain restrictions in the Arrow structure

The standard illustration of the difficulties that can arise in Arrow's structure is the Condorcet voting paradox. Table 12.1 shows two profiles of preferences over three alternatives (election candidates, for example) for which following the majority on each pair-wise comparison leads to an intransitive outcome.

	Profile P			Profile P'		
individual	1st	2nd	3rd	1st	2nd	3rd
1	a	b	c	a	c	b
2	b	c	a	c	b	a
3	c	a	b	b	a	c

Table 12.1: Voting paradox

Any domain restriction that excludes the possibility that an individual states one of the preferences in each of the profiles in table 12.1 is sufficient to avoid intransitivity over $\{a, b, c\}$ however many individuals state each of the other preferences (Sen & Pattanaik, 1969). The possibilities are given in table 12.4.

	excludes from		
restriction	profile P	profile P'	description
a not first	1	1	single-troughed
a not middle	3	3	
a not last	2	2	single-peaked
b not first	2	3	single-troughed
b not middle	1	2	
b not last	3	1	single-peaked
c not first	3	2	single-troughed
c not middle	2	1	
c not last	1	3	single-peaked

Table 12.2: Domain restrictions

When there are more than three alternatives, majority voting gives a transitive outcome provided that one of these restrictions holds for each triple of alternatives.

Figure 12.1 shows that 8 orderings of 4 alternatives remain if preferences over each triple are single-peaked. Specifically, b is not last in triples $\langle a, b, c\rangle$ and $\langle a, b, d\rangle$ and c is not last in triples $\langle a, c, d\rangle$ and $\langle b, c, d\rangle$.

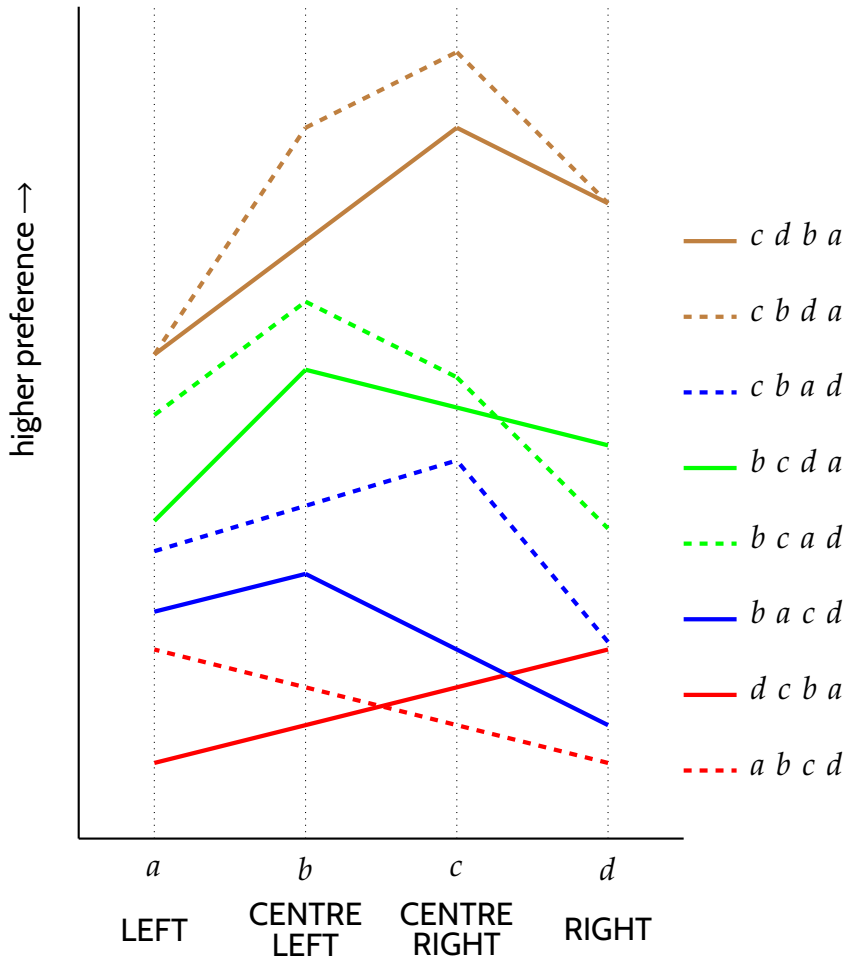

Figure 12.1: Four alternatives: single-peaked preferences

A common motivation for considering that preferences are single-peaked is to assert a meta-agreement[53] that political parties can be placed on a left-right political spectrum, and that voters have views that conform to those placings in the manner illustrated in figure 12.1. No voter prefers both 'extremes' to centre parties.

[53] List (2002) refers to 'agreements at a meta-level' to contrast with agreements at a substantive level. At the meta-level, people agree about the extent to which they can disagree.

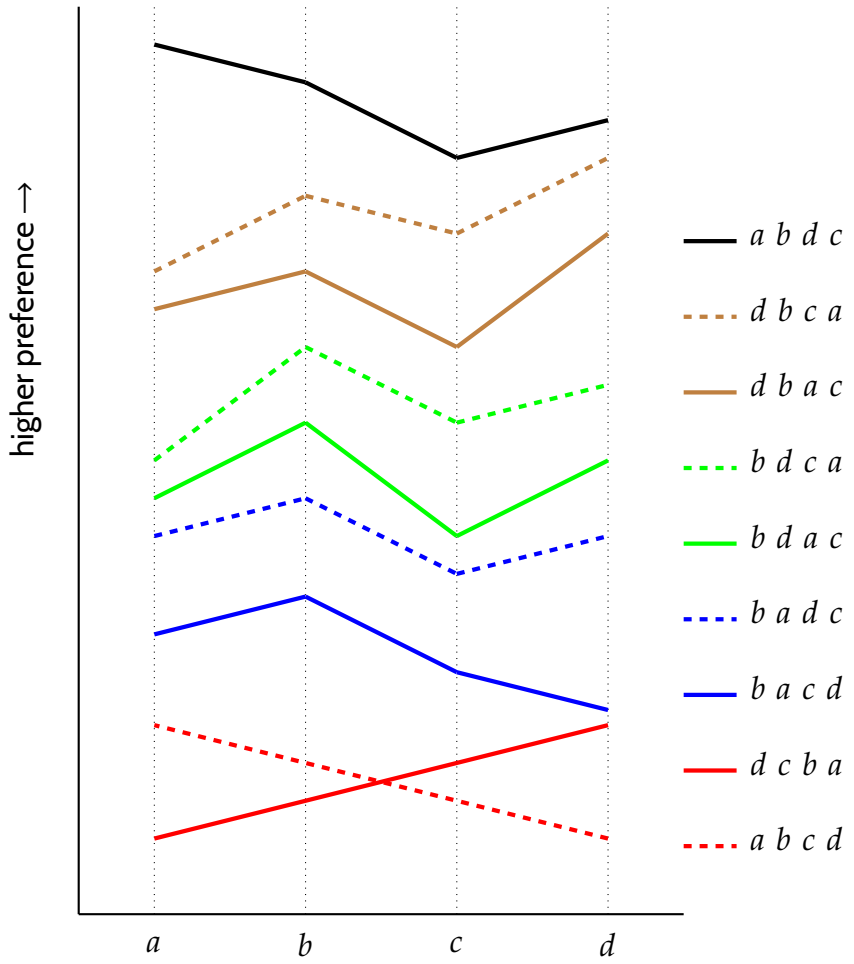

Figure 12.2: Four alternatives: alternate peaked and troughed orderings

The LEFT-RIGHT motivation is not applicable to the domain restriction illustrated in figure 12.2. In this figure, b is not last in triples $\langle a, b, c\rangle$ and $\langle a, b, d\rangle$, and c is not first in $\langle a, c, d\rangle$ and $\langle b, c, d\rangle$. This alternation of peaked and troughed orderings allows 9 distinct individual orderings compared to 8 in the single-peaked example of figure 12.1. Single-peakedness has some intuitive advantages when a LEFT-RIGHT spectrum or other meta-agreement is possible, but other domain restrictions are possible that allow numerically more possible orderings for any $n \geq 4$ (Craven, 1996). Fishburn (2002) more fully surveys the literature on domain restrictions in the Arrow structure.

12.3 The Mirkin structure

The structure established by Mirkin (1975) involves the aggregation of a profile of m partitions of S or equivalence relations over S, one stated by each individual. Each relation defines $r \geq 2$ equivalence classes in S, some of which may be empty. The aggregator determines an outcome which is a single equivalence relation over S that also defines r classes, some of which may be empty. So the domain of the aggregator is the set of all partitions of S into at least 2 and no more than r, non-empty classes.

The conditions (or, in his terminology, axioms) used by Mirkin are close parallels to those used by Arrow. The domain of the aggregator is unrestricted (given r). The aggregator satisfies an independence condition that the outcome for the co-location in a class of any pair of objects depends only on individual statements about the co-location of that pair of objects. In parallel terminology to that used here the independence condition is one of ***pair consistency***. The aggregator satisfies a binary non-imposition condition that for each pair of objects there is some profile for which that pair is co-located in the outcome and some other profile for which that pair is not co-located in the outcome. The aggregator also satisfies a ***pair monotonicity*** (or positive association) condition that additional support for the co-location of a pair does not lead to a perverse change in the outcome. These conditions together imply a ***pair unanimity*** condition that (i) if every individual co-locates a pair of objects, then that pair is co-located in the outcome and (ii) if no individual co-locates a pair of objects, then that pair is not co-located in the outcome.

Mirkin's statement of conditions parallels that originally used by Arrow, where the unanimity condition followed from non-imposition and monotonicity conditions. This structure is also binary because the evidence and outcome are binary (equivalence) relations and the conditions are expressed in terms of components of the binary relation.

Mirkin offers three possible contexts for which his structure is relevant. Paraphrased, these are

1. **Group choice** n individuals express their judgments about a set of objects, and the objective is to find out the group judgment on the basis of individual judgments. The judgment of each individual is characterised by a partition into groups of sufficiently similar objects. The outcome partition is then the group judgment.

2. **Research classifications** Social survey data on a single population is considered by various researchers from different aspects, so that every researcher has a classification for the population. Is it possible to find a representative classification?

3. **Factor analysis** of nominal attributes. Suppose there are n nominal attributes [properties in our terminology] describing a surveyed human population. The attributes are assumed to be external manifestations of some latent factor. Nominal attributes are determined by corresponding partitions into groups of persons with the same values, finding this "latent factor" involves aggregation in Mirkin's structure.

Each of these suggested contexts recognises that the evidence and the outcome in Mirkin's binary structure are in the form of partitions or equivalence relations. It is possible to say which pairs of objects are co-located by each individual and in the outcome, but not to say in which category each object is located. In the musical genre context, this places emphasis on whether Rosa and her influencers regard compositions as sufficiently similar that they are co-located (with due regard to the need for transitivity), and not on the genre location of each composition.

The parallel to Condorcet's voting paradox in Mirkin's structure is given in table 12.3.

individual	Profile P	Profile P'
1	$[\{a, b\}, \{c$, all others$\}]$	$[\{a, b\}, \{c$, all others$\}]$
2	$[\{a, b, c\}, \{$all others$\}]$	$[\{a, c\}, \{b$, all others$\}]$
3	$[\{b, c\}, \{a$, all others$\}]$	$[\{b, c\}, \{a$, all others$\}]$

Table 12.3: Profiles of equivalence classes

With profile P, a majority co-locates $a\&b$, a majority co-locates $b\&c$ but a minority co-locates $a\&c$ yielding an intransitivity. With profile P', a minority co-locates each pair, and so a, b, c must all be in different classes in the outcome, requiring three classes when $r = 2$.

12.3.1 Domain restrictions in the Mirkin structure

Both of the profiles of table 12.3 are excluded by a domain restriction that prevents any individual from co-locating one pair in the triple - say $b\&c$. Then there remain available only two equivalence relations that define classes $[\{a, b\}, \{c\}]$ and $[\{a, c\}, \{b\}]$. The outcome can be determined by following the majority.

If $n \geq 4$ the same consideration applies to each triple, and the domain restriction must exclude the co-location of additional pairs. For example, if $n = 4$ and $r = 2$:

- the domain restriction that no individual co-locates $b\&c$ and no individual co-locates $a\&c$ leaves two equivalence relations that define classes $[\{a, b, d\}, \{c\}]$ and $[\{a, b\}, \{c, d\}]$;
- the domain restriction that no individual co-locates $b\&c$ and no individual co-locates $a\&d$ leaves two equivalence relations that define classes $[\{a, b\}, \{c, d\}]$ and $[\{a, c\}, \{b, d\}]$.

In each case, the domain restriction that enables the use of a majority aggregator removes all but two equivalence relations from all profiles.

12.4 The Maniquet & Mongin structure

The other main focus in the literature is the structure set out by Maniquet & Mongin (MM). Each individual states a classification of S into r named categories. The only restriction on the individual classifications is that no category is empty. The aggregator defines an outcome which is also a classification of S into r non-empty categories.

The requirement that all categories in the outcome are non-empty implies that the aggregator is **surjective**. This requirement is important because it is the way in which the outcome locations of each object relate to one another - just as transitivity is required in relation to the outcomes for pairs of objects in the binary structures of Arrow and Mirkin, and in the structure introduced here in chapter 4. Plainly, any majority-based aggregator can fail to be surjective because a category can be empty in the outcome (profile P' in table 12.3) if there is some, but not majority, support for the inclusion of any object in a category.

12.4.1 Example contexts

MM give several examples, emphasising the role of the condition that no category is empty.

- If a group of law-makers is conferring legal status on a social category, MM argue that they will do so only if they believe that each category is applicable to some citizens. In a country "where democratic principles hold", one may further expect that the categories have been agreed on between the lawmakers and the citizens prior to being used in practice. Accordingly, citizens would not leave any category unfilled, even though they can disagree on its precise extension. MM report that the group identification literature alludes to political examples that seem to warrant this analysis, including the consideration of Israeli citizenship by Kasher & Rubinstein (1997).

- A panel of astronomers meets to classify distant celestial bodies into stars, exoplanets, brown dwarfs …. Each astronomer proposes his own classification, and the chair tries to turn these individual data into an authoritative classification. The classification is well-established on prior grounds, so if the set of celestial bodies under consideration is large enough, neither the individual astronomers nor the chair will leave any of the four categories empty.

- MM use an example from Dokow & Holzman (2010) in which

  > "…[there is] a conservation camp, in which each enrollee [worker] has one and only one activity, like planting a tree or opening a trail, such activities typically require enrollees to work side by side, and the camp wants each of them to be carried out. There are several supervisors, and each proposes an assignment [of tasks to the workers. The ambition of the camp manager is] to define a collective task assignment from these individual proposals." (page 9)

  Each supervisor states an assignment of the workers according to the supervisor's view of their skills, given the range of tasks to be accomplished, and the assumption that no category is empty follows if the supervisors and the manager want to ensure that every task is covered by the assignment of at least one worker to it.

The MM structure could be used in the musical genre context as an alternative to that introduced in chapter 4. The domain of the problem is different because the MM structure requires that no genre is empty in any individual (influencer) classification, and similarly in the outcome (Rosa's) classification. Empty categories are allowed in chapter 4 - and the conditions on the aggregator are different in the two structures. However, as section 12.5 shows, in many circumstances the MM conditions entail those used in chapter 4, and the MM structure is unambiguously more restrictive.

12.4.2 The MM conditions

MM demonstrate that there must be a dictator (or dominant individual) if[54]

- every individual states a classification into $r \geq 2$ non-empty categories
- the domain of the aggregator is unrestricted so that every object is eligible for every category;
- the aggregator satisfies ***class independence*** so that the outcome location of any object depends only on the stated locations of that object by the individuals;
- the aggregator satisfies ***class unanimity part 1*** that the outcome location of x is C_s if every individual locates x in C_s;
- the aggregator is **surjective** which requires that no category is empty in the outcome.

Note that ***class unanimity part 1*** differs from our ***class unanimity*** condition because the latter (section 10.1.3, page 151) requires the stronger condition that x is not in category C_i in the outcome if no individual locates x in C_i.

The MM structure is **unary** because the evidence and outcome are locations of objects in categories and the conditions are expressed in terms of those unary locations. No binary relation is involved in this structure.

A parallel example to Condorcet's voting paradox in MM's structure is given in table 12.4.

[54]The two-category case is a theorem from Kasher & Rubinstein (1997). The notation is that of chapter 10.

individual	
1	$C_1 = \{a, b, \text{others}\}, C_2 = \{c\}$
2	$C_1 = \{a, c, \text{others}\}, C_2 = \{b\}$
3	$C_1 = \{b, c, \text{others}\}, C_2 = \{a\}$

Table 12.4: Profile of classifications

The profile of table 12.4 shows that a majority aggregator of the unary statements cannot be used to determine the outcome in the structure of MM. A majority locates every object in C_1, leaving C_2 empty.

The surjectivity requirement that no category is empty in the outcome is crucial: if empty categories were allowed, a majority aggregator can be used when $r = 2$ and predominance aggregators when $r > 2$. Both satisfy the unary independence and unanimity conditions used by MM but not their surjectivity requirement.

The simplest way to avoid the dictatorial conclusion in the MM structure is to assume a domain restriction that there is an archetype in each category. Then the MM unanimity condition entails that no category is empty in the outcome, and the outcome location of any other object can be determined by following the majority or predominance depending on the categories for which that object is eligible.

12.5 Comparing MM and the structure of chapter 4

The structure that is introduced in chapter 4 and examined more formally in chapter 10 involves both unary and binary conditions on the aggregator, and does not include the surjectivity requirement. ***Class unanimity*** (both parts) and ***class independence*** are unary; ***pair unanimity*** is binary. This hybrid structure is possible because the unary statements of the form "x is in category C" together define an equivalence relation that allows the meaningful use of both unary and binary conditions. Furthermore, as theorem 1 shows, a majority aggregator can be used provided that no object is eligible for more than two categories without requiring

that there is an archetype in each category. So the domain restrictions that avoid a dominant individual differ from those needed in the MM structure.

However, when the domain is unrestricted, there are three or more categories and more objects than categories, the MM conditions (surjectivity, ***class unanimity*** part 1 and ***class consistency***) entail the conditions used in theorem 2, so that the MM result is entailed by that theorem.

Using the notation of chapters 10 and 11, also define a **cover**[55] for Γ (the set of r categories) which is a subset $T \subset S$ that contains exactly $r = \sharp\Gamma$ objects. Then profile P contains a cover T for Γ just if each individual locates exactly one object from T in each category, and so no individual leaves any category empty, and no individual co-locates any two objects from T: $\forall i \in I\ \forall C_s \in \Gamma\ \exists x \in T\ P_{ix} = C_s$. If the domain is unrestricted, it is always possible to find a profile that includes a cover for Γ.

Class unanimity part 2 is the condition that x is not in C_s in the outcome if no individual locates x in C_s

Result 24. *When the domain of f is unrestricted, the MM conditions entail* **class unanimity** *part 2.*

This is lemma 1 in Maniquet & Mongin.

Proof. Suppose that the result is false, Then $\exists P \in \Pi\ x \in S, C_s \in \Gamma\ [\forall i \in I\ P_{ix} \neq C_s\ \&\ q_x = f_x(P) = C_s$. then define $P' \in \Pi$ such that

$\forall i \in I\ P'_{ix} = P_{ix}$ and

for some $y \in S\ y \neq x\ \forall i \in I\ P_{iy} = C_s$ and

$T = \{x, y, w_1 \ldots w_{r-2}\}$ is a cover for Γ in P' and

[55] This terminology is not connected to the use of 'cover' in topology.

$\forall z \notin T\ P'_{iz} = C_s$.

Then ***class consistency*** $\rightarrow q'_x = q_x = C_s$, ***class unanimity*** part 1 $\rightarrow q'_y = C_s$ and $\forall z \notin T\ q_z = C_s$. There remain $r-2$ objects in S but surjectivity requires that the remaining $r-1$ categories be non-empty, which is not possible. So the result is not false. □

Pair unanimity also consists of two parts. Part 1 requires that $x\&y$ are co-located in the outcome if every individual co-locates $x\&y$. Part 2 requires that $x\&y$ are not co-located in the outcome if no individual co-locates $x\&y$.

Result 25. *When the domain of f is unrestricted, the MM conditions entail* **pair unanimity** *part 2.*

Proof. Suppose that ***pair unanimity*** part 2 does not hold. Then $\exists P \in \Pi\ x, y \in S\ C_s \in \Gamma\ [\forall i \in I\ P_{ix} \neq P_{iy}\ \&\ f_x(P) = f_y(P) = C_s]$. Then define $P' \in \Pi$ such that $[\forall i \in I\ P'_{ix} = P_{ix}\ \&\ P'_{iy} = P_{iy}]$, $W = \{x, y, w_1 \ldots w_{r-2}\}$ is a cover for Γ in P' and $\forall z \notin W\ [\forall i \in I\ P'_{iz} = C_s]$. Then ***class unanimity*** part 1 and ***class consistency*** $\rightarrow \forall z \notin W\ f_z(P') = f_x(P') = f_y(H') = C_s$ and the remaining $r-2$ objects in W are insufficient to ensure that all the remaining $r-1$ categories are non-empty in the outcome $f(P)$ contrary to the surjectivity requirement. So ***pair unanimity***(b) holds. □

Result 26. *When the domain of f is unrestricted and $n > r$, the MM conditions entail* **pair unanimity** *part 1.*

Proof. Suppose that ***pair unanimity*** part 1 does not hold. Then $\exists P \in \Pi\ x, y \in S\ C_s, C_t \in \Gamma\ [\forall i \in I\ P_{ix} = P_{iy}\ \&\ f_x(P) = C_s \neq f_y(P) = C_t]$. Define $P' \in \Pi$ such that $[\forall i \in I\ P'_{ix} = P_{ix}\ \&\ P'_{iy} = P_{iy}]$, $W = \{x, w_1 \ldots w_{r-1}\}$, $y \notin W$ is a cover for Γ in P' and $\forall z \notin W \cup \{y\}\ [\forall i \in I\ P'_{iz} = C_s]$. Note that this is possible only if $n \geq \sharp(\{y\} \cup W) > r$.

Class independence $\rightarrow f_x(P') = C_s \;\&\; f_y(P') = C_t$. ***Pair unanimity*** part 2 $\rightarrow$ $[\forall u, v \in W \; f_u(P') \neq f_v(P')]$. By hypothesis $W' = \{y, w_1 \ldots w_{r-1}\}$ is also a cover for Γ in P'. But $x \in W \;\&\; y \in W' \;\&\; f_x(P') \neq f_y(P')$. So $\forall u, v \in W \cup W' \; f_u(P') \neq f_v(P')]$. $\sharp W \cup W' = r + 1$ and $\forall u, v \in W \cup W' \; f_u(P') \neq f_v(p')$. This requires $r + 1$ distinct categories. So ***pair unanimity*** part 1 holds. □

Theorem 2 is valid in the hybrid structure when there are more than two categories and the domain is unrestricted (so that $\forall s \;\; E_s = S$). However, unlike the structure of MM, this allows that some categories can be empty in any individual classification. The two structures have different domains even when there are no domain restrictions entailed by eligibilities.

The results needed to prove theorem 2 require the choice of profiles that are consistent with the restrictions given in tables 10.2 to 10.8. Provided that $n > r$ there are sufficient additional objects to form a cover for Γ that is consistent with the restrictions in each table. So the tables can be adapted to require only profiles in the domain used in the MM structure. The choice of a profile in which other objects help to form a cover for Γ does not affect the proofs because ***class consistency*** rules out any impact from the locations of those other objects on the objects involved directly in the proofs.

For example, the profile used to demonstrate result 6 on page 156 is restricted as in column 10.2d of table 10.2. If there are four categories, the profile of table satisfies these restrictions and $\{w, x, y, z\}$ forms a cover for Γ.

individuals	P_{iw}	P_{ix}	P_{iy}	P_{iz}
$i \in D$	C_t	C_s	C_u	C_v
$i \in D^c$	C_v	C_t	C_s	C_u

Table 12.5: A cover for result 6

There must be at least one more object than categories to enable the definition of a cover for profiles that are consistent with column 10.2b of table 10.2 (and for

scenario 2 in column 10.7a of table 10.7) because in those profiles two objects are co-located by some individuals and so cannot both be in a cover for Γ.

With this modification of the proof, theorem 2 applies to the domain used in the MM structure. Results 24 to 26 show that the MM conditions entail the conditions used in the hybrid structure and so theorem 2 entails the MM dictatorship (dominant individual) result provided that $n > r > 2$.

12.5.1 Interpretation

Social choice theory rejoices in technical details, and the formal result above applies only when there are more than two categories and more objects than categories. Both the MM theorem and our theorem 2 apply when there is the same number of categories as objects, but this possibility is not included in result 26. The reason that theorem 2 cannot be applied in this way to the MM structure is that if there is the same number of objects as categories, and no category can be empty in any individual classification, then S is a cover for Γ in every profile. ***Pair unanimity*** part 1 applies to no pair of objects, and part 2 applies to every pair making the proof strategy untenable when the domain changes from that of the hybrid structure to that of the MM structure.

In a 'practical' application such as the task assignment challenge on page 190 there can be legitimate interest in a situation in which there is the same number of workers as activities. In the MM structure, every supervisor is constrained to locate exactly one worker to each task, and the surjectivity condition requires that all activities can be performed. In the hybrid structure supervisors can leave activities without a nominated worker, and some activities might not be performed by any worker.

But in the language example, there is little advantage in considering a context in which there is the same number of genres as compositions. Then communication is not much enhanced by the introduction of the genre terms - it can happen

as easily by referring to compositions directly. Genres are intended as collective terms for several - perhaps many - compositions. Communication is then simplified by applying the collective term rather than listing the compositions on each occasion. The challenge for Rosa, or for a group of people seeking to devise a common terminology, is to devise the collective terms - that is the genre classifications - given the diversity of views of the influencers.

12.6 Summary

This chapter does not add anything directly to the central topic of concept formation in the presence of conflicting evidence. It considers instead how the structure of chapters 4 and 10 relates to other structures considered in the social choice literature.

That literature includes discussion of domain restrictions that allow the consistent use of majority methodologies the aggregation of orderings in Arrow's structure. We extend that to show a simple restriction that applies in the aggregation of equivalence relations in Mirkin's structure.

The structure introduced by Maniquet & Mongin aggregates classifications and is shown to be a special case of the structure of chapter 4: their conditions entail all of the deference principles used in chapter 4 if, as they assume, the domain is unrestricted. But their more stringent conditions do not allow the consistent use of a majority methodology when the domain is simple.

Chapter 13

Conclusion

Disagreement is widespread in any social organisation in which practices and policies are based on individuals' statements. One focus within philosophy has been a discussion of how disagreement can arise, or be explained, when there is a fundamental expectation that propositions are either true or false, so that disagreement can arise if some individuals make a mistake by asserting a false proposition or if they are presenting different propositions, plausibly keeping implicit the source of their differences.

Different uses of concepts arise from the recognition that there are propositions that are important to a language learner but for which it is not possible to discern truth or falsity. This failure of discernment can arise for reasons that include the incompleteness of knowledge that might one day be resolved (whether joint pain is an instance of ARTHRITIS or of LUPUS), different evaluations of evidence (whether *Rhapsody in Blue* instances JAZZ or CLASSICAL) or vagueness (fundamental lack of clarity about the placing of the boundary between dishes that instance HOT or MILD). Coping with this epistemic problem - what we cannot or do not know - is fundamental to the use of language in communication.

Our fundamental question (page 20) has been

Is it possible to find a satisfactory compromise when there is disagreement or diverse evidence about extensions of concepts?

The generic answer that we have reached is that extensive disagreement cannot be reconciled in a wholly satisfactory way. The nature of 'satisfactorily' depends to some extent on the context, but generally includes an ambition to ensure that the attribution of objects to concepts is consistent with the conventions of the language, and is done in a way that allows piecemeal decisions rather than requiring a simultaneous consideration of many objects. In addition, there are circumstances in which there is no conflict of evidence - and in those circumstances it would seem perverse to reach a conclusion that conflicts with all the evidence. When the extent of potential disagreement is sufficiently extensive, failure to reconcile the diverse evidence is generally manifest in a conclusion that the ambition to meet these criteria entails that only one source of evidence can be considered, even though there are several available.

This single source of evidence might be one of the language learner's influencers (chapters 4 and 7), a single property of the objects in question (chapter 6), her own internal evaluation which cannot be impacted by the views of others (section 6.6) or, in appropriate contexts, an individual's self-designation within a category or concept (chapter 9). A possible 'escape route' from this conclusion might be found when the language learner is using evidence from the properties of the objects, when she might additionally consider for herself the intensity with which she thinks properties support concept locations. Deriving meaningful comparable evidence of the intensity with which her influencers support particular locations is much more problematic - comparable with the difficulties of making inter-personal comparisons of utility, and hence of identifying the 'greatest good of the greatest number' in ethics or politics.

When disagreement about each issue is limited to only two possibilities, it is usually possible to avoid the dominant outcome if the language learner follows the

majority of the evidence. In the main result (theorems 1 and 2) that is applicable to the musical genres context, there is a straightforward contrast between circumstances in which the conventions of the language entail that there are no more than two possibilities for the location of each composition and circumstances in which one or more is not so limited. This provides a clear demarcation in the extent of disagreement that can be satisfactorily reconciled. A parallel result (theorems 3 and 4) applies to the location of boundaries in the graded 'spicy dishes' context. The possibility of using a majority or preponderance methodology in a self- and other-designation context depends on whether it is reasonable to leave a concept empty even though someone self-designates in it.

Vagueness (chapter 8) is one possible source of different statements about the truth of propositions, and the conclusions that we have reached apply then too. Discussions of vagueness often involve the recognition that there are several possible precisifications of the concept(s) involved. A more fundamental issue arises when we allow that the language itself can be vague, so that each of the learner's influencers has his own idiolect and so his own set of eligibilities, with no evidenced argument to support the selection of one rather than another. Then the learner has two challenges whose resolutions can conflict if the scope for disagreement between influencers is sufficiently large. The learner defers to her influencers to determine her own idiolect, and then defers again to establish how she locates objects into concepts within that idiolect. With sufficient disagreement, it is possible that the learner ends up in a situation in which she tries to locate an object in a concept for which it is not eligible within her idiolect. Once again, too much scope for disagreement can entail an unsatisfactory outcome.

A significant insight here is that social choice theory can contribute to an understanding of language learning. But the analysis has taken us beyond considerations that are likely to be relevant in the examination of concept formation by a learner. Some of our results, particularly the conclusions about the outcomes from alternative structures in chapter 12 and the considerations of designation in chap-

ter 9, are more generally conclusions of social choice theory, and this dissertation extends some of the conclusions in the literature. Many conclusions in the social choice literature focus on contexts in which the domain is unrestricted, whereas the essence of a conventional language is likely to be the recognition that some theoretically possible usages of concepts do not conform to the eligibilities of the language. They constitute mistakes, and avoiding mistakes restricts the domain of the aggregation problem as it might be applied in many contexts.

Bibliography

[1] Adorno, T.W. (1939, 1989) On jazz. *Discourse* 12, 45-69. Translated by Jamie Owen Daniel. *Über Jazz* was first published in 1936 under the pseudonym Hektor Rottweiler.

[2] Alcantud, J.C.R. & Laruelle, A. (2020) Independent collective identity functions as voting rules. *Theory and Decision*, 89, 107-119.

[3] Arrow, K.J. (1951) *Social choice and individual values*. Cowles Foundation Monograph, 12. (New York: John Wiley).

[4] Arrow, K.J. & Hahn, F.H. (1971) *General competitive analysis*. (San Francisco: Holden-Day).

[5] Belleri, D. (2010) Relative Truth, Lost Disagreement and Invariantism on Predicates of Personal Taste. in Crespo, M.I., Dimitris Gakis, D. & Weidman-Sassoon, G. (eds.) *Proceedings of the Amsterdam Graduate Philosophy Conference*. Department of Philosophy, Universiteit van Amsterdam.

[6] Belleri, D. & Palmira, M. (2013) Towards a unified notion of disagreement. *Grazer Philosophische Studien*, 88, 139-59.

[7] Black, D. (1948) On the rationale of group decision-making. *Journal of Political Economy*, 56, 23-34.

[8] Burge, T. (1979), Individualism and the mental. *Midwest Studies in Philosophy*, 73-121.

[9] Burge, T. (1986), Intellectual norms and foundations of mind. *The Journal of Philosophy*, 83, 697-720.

[10] *The Cambridge Dictionary* https://dictionary.cambridge.org/dictionary/english/

[11] Chilliworld. *The Scoville heat scale*. https://www.chilliworld.com/factfile/scoville-scale

[12] Chichilnisky, G. & Heal, G. (1983) Necessary and sufficient conditions for a resolution of the social choice paradox. *Journal of Economic Theory* 31, 68-87.

[13] Cho, W.J. & Ju, B-G. (2017) Multinary group identification. *Theoretical Economics*, 12, 513-531.

[14] Condorcet, M. d. (1785) *Essai sur l'application de l'analyse à la probabilité des décisions rendues a la pluralité des voix*. (Paris).

[15] Craven, J. (1992). *Social Choice*. Cambridge: Cambridge University Press.

[16] Craven, J. (1996) Majority-consistent preference orderings. *Social Choice and Welfare*, 13: 259-67.

[17] Craven, J. (2022) Self-designation and group allocation. *Theory & Decision*. https://doi.org/10.1007/s11238-022-09882-z

[18] Craven, J. (2023) Domain restrictions in the aggregation of classifications. *Global Philosophy*. https://doi.org/10.1007/s10516-023-09670-6

[19] Debreu, G. (1959) *Theory of value*. Cowles Foundation Monograph, 17. (New Haven: Yale University Press).

[20] Dokow, E. & Holzman, R. (2010) Aggregation of non-binary evaluations. *Advances in Applied Mathematics*, 45, 487–504.

[21] Dummett, M. (1975). Wang's paradox. *Synthese* 30, 301-324.

[22] Fioravanti, F & Tohmé, F. (2020) Alternative axioms in group identification Problems. *Journal of Classification*. https://doi.org/10.1007/s00357-020-09378-x

[23] Fishburn, P. (2002) Acyclic sets of linear orders: a progress report. *Social Choice and Welfare*, 19, 431-447.

[24] Fodor, J.A. (1994) *The elm and the expert*. (Cambridge, Mass: The MIT Press).

[25] Frances, B. and Matheson, J. (2019), Disagreement. *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2019/entries/disagreement/

[26] Galili, L. (2020) *The other tribe: Israel's Russian-speaking community and how it is changing the country*. (Brookings). https://www.brookings.edu/research/the-other-tribe-israels-russian-speaking-community-and-how-it-is-changing-the-country/

[27] Gioia, T. (1998) *The history of jazz*. (Oxford: Oxford University Press).

[28] Goffin, R. (1946) *Jazz: from Congo to swing*. (London: Musicians Press).

[29] Greenberg, M. (2011) Legislation as communication? Legal interpretation and the study of linguistic communication. Chapter 10 in Marmor, A. & Soames S. *Philosophical foundations of language in the law*. (Oxford: Oxford University Press).

[30] Greenberg, M. (2014) Troubles for content I. Chapter 5 in Burgess, A. & Sherman, B. (ed) *Metasemantics: new essays on the foundation of meaning*. (Oxford: Oxford University Press).

[31] Gutman, P. (2003) Classical notes. http://www.classicalnotes.net/classics/gershwin.html

[32] Hare, R.M. (1952) *The language of morals*. (Oxford: Oxford University Press).

[33] Hobbes, T. (1651) *Leviathan*.

[34] Hume, D.(1748) *Philosophical essays concerning human understanding*. (London: A. Millar).

[35] Houy, N. (2007) "I want to be a J!": liberalism in group identification problems. *Mathematical Social Sciences*, 54, 59–70.

[36] Biography.com (2014) *Jelly Roll Morton, a biography*. https://www.biography.com/musician/jelly-roll-morton

[37] The Jewish Agency for Israel. (2019) *The law of return*. http://archive.jewishagency.org/first-steps/program/5131

[38] Kasher, A. (1985), Justice and affirmative action: naturalization and the law of return. *Israel Yearbook of Human Rights*, 15, 101-12.

[39] Kasher, A. (1993) Jewish collective identity, in *Jewish Identity* edited by Goldberg, D.T. & Krausz, M. (Philadelphia: Temple University Press).

[40] Kasher, A. & Rubinstein, A. (1997) On the question "Who is a J?", a social choice approach. *Logique & Analyse*, 160, 385-395.

[41] Kölbel, M. Faultless Disagreement. *Proceedings of the Aristotelian Society*, New Series, 104, 53-73.

[42] Kuhn, T. (1969) *The structure of scientific revolutions*. (Chicago: University of Chicago Press).

[43] Kuhn, T. (1977): Objectivity, value judgment, and theory choice, in *The essential tension*. (Chicago: University of Chicago Press).

[44] Lewis, D. J. (1965). *Introduction to algebra*. (New York: Harper & Row).

[45] Lewis, D.K. (1969) *Convention*. (Cambridge: Harvard University Press).

[46] List, C. (2002). Two concepts of agreement. *The Good Society*, 11, 72-9.

[47] List, C. (2011) Group communication and the transformation of judgments: an impossibility result. *Journal of Political Philosophy*, 19, 1-27.

[48] Maniquet, F. & Mongin P. (2016) A theorem on aggregating classifications. *Mathematical Social Sciences*, 79, 6–10.

[49] May, K.O. (1952) Intransitivity, utility, and the aggregation of preference patterns. *Econometrica*, 22, 1-18.

[50] Miller, A.D. (2008) Group identification. *Games and Economic Behavior*, 63, 188–202.

[51] Mirkin, B.G. (1975) On the problem of reconciling partitions, in Blalock, H.M., Aganbegian, A., Borodkin, F., Boudon, R., Capecchi, V. (Eds.), *Quantitative Sociology. International Perspectives on Mathematical and Statistical Modelling*, 441-9. (New York: Academic Press).

[52] Morreau, M. (2015) Theory choice and social choice: Kuhn vindicated. *Mind*, 124, 239-62.

[53] Nanry, C. (1979), *The jazz text*. (New York: Van Nostrand).

[54] Nanson, E.J. (1882) Methods of election. *Transactions and Proceedings of the Royal Society of Victoria*, 19, 197-240.

[55] Okasha, S. (2011). Theory choice and social choice: Kuhn versus Arrow. *Mind*, 120, 83-115.

[56] Okasha, S. (2015). On Arrow's theorem and scientific rationality: reply to Morreau and Stegenga. *Mind*, 124, 279-94.

[57] Panassié, H. (1942), *The real jazz*. Translated by Anne Sorrelle Williams. (New York: Smith & Durrell).

[58] Plunkett, D. & Sundell, T. (2013) Disagreement and the semantics of normative and evaluative Terms. *Philosophers' Imprint*, 13, http://hdl.handle.net/2027/spo.3521354.0013.023

[59] Reeves, P. (2011) Russian influx splits Israel over Jewish identity. *Independent*, https://www.independent.co.uk/news/world/middle-east/russian-influx-splits-israel-over-jewish-identity-699676.html

[60] Rescorla, M, Convention. *The Stanford Encyclopedia of Philosophy (Summer 2019 Edition)*, Edward N. Zalta (ed.). https://plato.stanford.edu/archives/sum2019/entries/convention/.

[61] Rubinstein, A. & Fishburn, P.C. (1986) Algebraic aggregation theory. *Journal of Economic Theory*, 38, 63-77.

[62] Sainsbury, R. M. & Tye, M. (2011) An originalist theory of concepts. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 85, 101-24.

[63] Sainsbury M. & Tye M.(2012) *Seven puzzles of thought and how to solve them: An originalist theory of concepts*. (Oxford: Oxford University Press).

[64] Samet, D. & Schmeidler, D. (2003) Between liberalism and democracy. *Journal of Economic Theory*, 110, 213-233.

[65] Sen, A.K. (1970) The impossibility of a Paretian liberal. *Journal of Political Economy*, 78, 152-157.

[66] Sen, A. K. (2017), *Collective choice and social welfare* (Extended edition). (London: Penguin).

[67] Sen A.K. & Pattanaik, P.K. (1969) Necessary and sufficient conditions for rational choice under majority decision. *Journal of Economic Theory*, 1, 178-202.

[68] Stegenga, J. (2015) Theory choice and social choice: Okasha versus Sen. *Mind*, 124, 263-77.

[69] Sung, S.-C. and Dimitrov, D (2005) On the axiomatic characterization of "Who is a J?", *Logique et Analyse*, 189/192, 101-112.

[70] Suzumura,K. (1981) On the Possibility of "Fair" Collective Choice Rule. *International Economic Review*, 22, 351-364.

[71] Williamson, T. (1994) *Vagueness*. (Abingdon, Oxon: Routledge).

[72] Williamson, T. (1999) On the structure of higher-order vagueness. *Mind*, 108, 127-143.

[73] Wilson, R. (1975) On the theory of aggregation. *Journal of Economic Theory*, 10, 89-99.

[74] Wright, C. (2001) On Being in a Quandary: Relativism, Vagueness, Logical Revisionism. *Mind*, 110, 45-98.

[75] Wright, C. (2006) Intuitionism, Realism, Relativism and Rhubarb, in Greenough, P. & Lynch, M.P. (eds.) *Truth and Realism* (Oxford, Oxford University Press).

Music

[76] C.P.E. Bach, *Magnificat*. Composed in 1749, Bach rewrote part of the score around 1779 during the transition from the baroque to classical eras.

[77] J.S Bach, *Air on a G-string* from BWV1068 (c1730) arranged by August Wilhelmj in 1871. Classic version as played (for example) by Anne-Sophie Mutter.

[78] L. van Beethoven, *Symphony No. 9 in D minor*, Op. 125. First performed in Vienna in 1824.

[79] L. Bernstein, *Prelude, Fugue and Riffs* composed in 1949. First performed on *The World of Jazz* in 1955.

[80] B. Britten, *Peter Grimes*. First performed in London in 1945.

[81] D. Ellington, *Take the "A" train*. Composed by Billy Strayhorn in 1939.

[82] G. Gershwin, *Rhapsody in Blue*. First performed at the Aeolian Hall, New York, in 1924.

[83] G.F. Handel, *Semele*. First presented in concert form at Covent Garden in 1744.

[84] King Cole Trio, arrangement of Rachmaninov, *Prelude in C♯ minor*. [89]. First recorded in 1944.

[85] R. Vaughan Williams, *The Lark Ascending*. Composed c1914. First performed at the Shirehampton Public Hall in 1920.

[86] The Modern Jazz Quartet & The Swingle Singers. Arrangement of J.S Bach, *Air for G-string*.

[87] L. Morey (lyrics) & F. Churchill (music), *Some Day My Prince will Come*. From the 1937 Disney animated film *Snow White and the Seven Dwarfs.*

[88] L. Morgan, *The Sidewinder*. Released in 1964 on the *Blue Note* label.

[89] S. Rachmaninoff, *Prelude in C♯ minor*, Op. 3. First performed in Moscow in 1892.