The question of form has firmly reasserted itself into musicological consciousness in the last
two decades. Renewed attempts have been made at understanding instrumental form, not
only from the safe distance of historical perspective (such as exploring the history of music
theory) but also in terms of new analytical approaches and – even in the age of postmodern
incredulity – grand unified systems. Of these, the Sonata Theory of James Hepokoski and
Warren Darcy (2006) is the most prominent example; its approach and terminology have
productively infiltrated much analytical writing in recent years but remain in some quarters
contentious. The challenge thrown down by this theory has since sustained several
conferences and publications, encouraging further and sometimes heated debate which does
not look like abating anytime soon.1 As one writer has noted, ‘It seems that nineteenth-
century Formenlehre has made a surprise comeback at the start of the twenty-first century’
(Riley 2008, p. 590).

Within the battle-lines drawn in this contested terrain, the nature of concerto form
remains an especially fraught issue. Debate has raged since the early nineteenth century and
the birth of Formenlehre over whether first-movement concerto form should be treated as
distinct from the sonata form or as merely a special case of it. This question is given greater
historical complexity in that as the nineteenth century progresses many prominent examples
in the genre seem to chart a movement towards the merging of the two. The present article
seeks to contribute towards current theorising of concerto form. It looks at two related works
from the 1830s that have notable complications for the models developed by Hepokoski and Darcy or, in their parlance, contain multiple ‘deformations’. These are, namely, Ignaz Moscheles’s Concerto No. 7 in C minor (*Pathétique*), Op. 93, dating from 1835–6 and published in 1838, and William Sterndale Bennett’s Concerto No. 4 in F minor, Op. 19 (1838), published in 1839.

From a twenty-first-century perspective, both Moscheles and Bennett must be reckoned as *Kleinmeister*. There has long been an assumption in modern, post-Kantian aesthetics that the works of such figures must necessarily provide the formal paradigms which great composers of genius break or transcend (or, in Sonata Theory terms, supply the defaults which are overridden). If one of the defining characteristics of great composers in the Romantic era is originality, and since originality is a relational property, some works must be providing the norms, and these will inevitably be those of second-tier figures. The two exist in true Master-Slave interdependency. Or considered historically, if works of genius are original but nonetheless exemplary and thus can provide a model to future generations – and (in a suitably modernist, Bloomian manner) if only the strongest artists are equipped with the capacity to misread and transcend the canonic examples of the past – then the definition of lesser artists is that they merely imitate past masters, rather than subvert them. Either way, the works of the *Kleinmeister* give us copies, whether of the past or of the present; they do not boldly set out the new path that leads to the future.

Such viewpoints have, admittedly, been less prevalent in the twenty-first century, and one of the implications of this paper is to call further into question the empirical veracity of the rather vulgar modernism underlying this belief. The two works examined here are in fact quite radical, but at the same time they engage dialogically not only with each other but also with earlier examples in the genre. Both concertos thus provide challenges to recent models
of form, affording us the opportunity to explore a host of issues concerning the theory of
concerto form and its relationship to the sonata.

This discussion may be further situated within ongoing attempts to deepen our
historical understanding of the development of the concerto genre. The work of Stephan
Lindeman in particular has opened awareness of the innovatory designs in the early
nineteenth-century concerto beyond the canonic Austro-German examples. More explicitly,
Julian Horton (2011, pp. 46–7) has recently offered cogent criticism of the Mozart/Viennese-
centric orientation of conventional concerto narratives. Instead, he calls for a ‘de-centred
theory of form, which allows localised repertories to dictate their own frames of reference and
which maps relations between them as instances of intertextuality’ (p. 82). Horton has gone
on to claim that

Intersections of [the Mozartian] paradigm with those developed for coeval histories
afford the opportunity to make broader theoretical assertions and trace patterns of
reception. Schumann’s response to Field is a significant example; Sterndale Bennett’s
marriage of the London model with Mozartian and Mendelssohnian influences
constitutes another fruitful line of enquiry. (Horton 2011, p. 82.)

In looking at two quite specific cases, this article responds to the same urge to
understand the nature of nineteenth-century concerto form more fully. To this end, it
examines the application of Sonata Theory to these two concertos by Moscheles and Bennett,
demonstrating the resourcefulness and flexibility of this method, but its also potential
difficulties, namely the primacy it affords to tonality in determining form at the expense of
parameters such as theme and texture. Although I focus primarily on formal issues and
competing theories of the concerto genre, the appeal of Sonata Theory to historical agency
will also necessitate recourse to biography and contemporary reception at later stages of the
discussion. Ultimately, I argue that although Sonata Theory is useful as both an analytical
and hermeneutic tool, there are still some issues that need to be resolved and much that
remains to be understood concerning the development of nineteenth-century concerto form.
Expositionary background

A standard, commonly accepted account of the development of first-movement concerto form across the first half of the nineteenth century is of the Classical, Mozartian merger of Baroque ritornello concerto form and sonata – what Hepokoski and Darcy call the Type 5 sonata – gradually giving way to a more purely sonata-orientated design (the Type 3 sonata), doing away with the clear distinctions between solo and ritornello sections (see Hepokoski and Darcy 2006, pp. 434–5). Fundamental to this progression is the suppression of the so-called double exposition in works from the 1820s and 30s onwards into a single exposition of thematic materials, combining soloist and orchestra in a unitary design. Although Lindeman (1999, p. 12) claims ‘the full transformation of the “first ritornello” into a “tutti exposition” by later composers to Mozart was (and continues to be) one of the most debated aspects of concerto form’, he agrees that by the years 1830-40 the first ritornello and first solo section (R1 and S1 in Hepokoski and Darcy’s terminology) were often being conflated into a unitary exposition (ibid., pp. 25–6). Moreover, as Jane Stevens has shown, by the second quarter of the nineteenth century theorists such as Marx and Czerny were increasingly reading the concerto as a subset of the sonata, rather than as arising independently and possibly slightly earlier, sharing at best common premises about general tonal behaviour (1974, pp. 50–2; see also Simon 1957). Typically, Mendelssohn’s mature concertos are seen as pivotal works here, though there are in fact earlier precedents, such as (in their different ways) Spohr’s Violin Concerto No. 8 in A minor ‘in modo di scena cantante’ (1816), Cramer’s Concerto No. 8 in D minor (c. 1819) and Weber’s Konzertstück Op. 79 (1821).

Both the works considered here are cast in a single exposition form, but both have further features that complicate and problematise type 5 and type 3 sonata forms – and potentially types 1 (sonatina, or sonata without development), 2 (binary) and 4 (sonata-rondo).
too. Both works are similarly unusual in that they return to the tonic minor for the ‘R2’ tutti statement which, by convention, signifies the closure of the ‘larger exposition’ (R1–S1–R2); and probably as a consequence of the reappearance here of their primary theme material in the tonic following the medial caesura, the recapitulation of the primary theme is drastically excised, complicating the post-expositional stages of the form. Of the two, the Bennett is the more radically deformed, but it appears to take its bearings substantially from the Moscheles.

Fig. 1a summarises the basic layout of Moscheles’s movement. In general thematic and textural layout, the movement conforms fairly closely to frameworks for first-movement concerto form set out by William Caplin and by Hepokoski and Darcy (the latter’s subtypes A, B, C and E, both given in Fig. 2). Two orchestral ‘ritornello’ sections (R1 & R2) flank the solo exposition (S1), followed by solo-dominated development section (S2), a brief (optional) tutti interjection (R3) as retransition (V of C minor, conforming here to Hepokoski and Darcy’s subtype A based on Koch’s 1793 concerto plan), a recapitulation centred on the soloist (S3) and a final tutti (R4, here, as customary for Moscheles, without cadenza). What is notable thematically is that the recapitulation proceeds, after the briefest four-bar prolongation of the dominant in dramatic diminished seventh chords in the piano, from the secondary theme (S1).

This design is not the most commonly encountered but is nevertheless accounted for within the context of Sonata Theory as a merging of the Type 5 (concerto) sonata with Type 2 (binary), or what Hepokoski and Darcy designate as (Type 5) concerto subtype E and name Mozart’s K. 491 in C minor as an example – a work which, as we shall see later, may have
influenced Moscheles’s concerto in other respects. However, when we look at the
movement’s tonal structure, the form of Moscheles’s work becomes far more unusual. What
is remarkable here is that the second tutti (R2) habitually used by Moscheles to demarcate the
end of the exposition opens firmly back in the tonic minor. On the face of it, Moscheles
gives us a sonata exposition that moves from C minor to the customary relative E<flat> major
and then back to C minor. Such a structure clearly contravenes what has been considered for
at least a century the most inviolate law of sonata construction: tonal opposition. The goal of
the exposition is none other than the same key and theme as heard at the outset.
This same conception is followed two years later in a concerto by William Sterndale
Bennett. Even if we were not to attribute so great a significance in itself to the mere structural
similarity between the two works, Bennett’s dedication of his concerto to Moscheles might
help persuade observers that it is very likely he was aware of the latter’s work (see Fig. 3).
Moscheles was at this time living in London and his concertos were widely played in the
British capital at the time the youthful Bennett was an up-and-coming pianist and composer at
the Royal Academy. Again (Fig. 1b), a minor-centred exposition moves to the expected
relative major for the second group, but this erstwhile harmonic region disintegrates back to
the tonic minor for the R2 passage. This time, however, the post-expositional stages are even
more problematised, a feature that will be addressed shortly.

Problems in formal definition
Before going any further, I would like to run through and counter a number of possible
objections to the apparently abnormal type of sonata design encountered in the two examples
here. Essentially, my argument is premised upon R2 being heard as part of the exposition, not the development. This reading is obviously couched in sonata-based terms; if the two movements were seen purely in terms of (Baroque) ritornello form, and therefore exempt from issues of sonata function, the question would not even arise. But to the extent that both the Moscheles and Bennett works feature a single exposition, and given the prevailing assumption of the concerto as tending towards sonata principles by this time, a sonata model seems entirely appropriate.

The earliest sonata-influenced accounts of concerto form, e.g. those by Kollmann and later by A. B. Marx) read the second orchestral ritornello as concluding the larger sonata exposition (Stevens 1974, pp. 36–56). Later English-language accounts tend to support the idea that R2 may have a dual function in concluding the exposition and initiating the development, reflecting the greater propensity for London-based composers to modulate in the second half R2 following an initial cadential affirmation (ibid., pp. 47–8, 56). This contention is supported by Julian Horton’s ‘English’ model for first-movement concerto form, ‘drawing together first-movement characteristics of the London concerto from Dussek through Cramer and Steibelt to Moscheles and Bennett’:

At the level of large-scale function, it comprises the seven-part R–S succession R1–S1–R2–S2–R3–S3–R4, overlaid onto a sonata form with double exposition (R1 and S1)... How we locate the start of the development depends upon the character of R2: if R2 is entirely post-cadential within the closing key of S1, then the development begins with S2; if R2 modulates immediately and without substantial post-cadential reinforcement of S1’s closing key, then R2 initiates the development; and if R2 divides into post-cadential and transitional segments, then the development properly begins where the R2 post-cadential material ends. (Horton 2011, p. 81.)

The only provision here for R2 as initiating the development (Horton’s second option) cannot apply to the two works in question, as in both the solo exposition does not close in the secondary tonality but veers back to the tonic minor for the tutti.

In their other examples in the genre, Moscheles’s and Bennett’s practice generally falls into Horton’s third category. In all six of Moscheles’s previous concertos, the
appearance of R2 serves as the goal of the solo exposition (S1), confirming the secondary
tonality in a strong orchestral statement before typically dissolving into a transition to the S2
‘development space’.9 This model holds even for the highly deformational Sixth Concerto
(Concerto fantastique), an example of a multi-functional cyclic work after the model of
Weber’s Konzertstück, J.B. Cramer’s Eighth Concerto and Alkan’s two Concerti da camera
Op. 10. Similarly, in Bennett’s three earlier published concertos, the R2 tutti statement
completes and closes the solo exposition with a return to Primary material in a closing theme
gesture, following an extended period of dominant preparation.10 Exactly the same procedure
occurs in the Fourth Concerto, the only difference being that the dominant in question has
changed from V/A<flat> (the ‘correct’ secondary tonality) at bars 207–216 to V/F minor (bars
221–234). This design – R2 as the rhetorical culmination of the exposition – is likewise
shared by a large number of late eighteenth- and early nineteenth-century concerto
movements. According to precedent, then, the listener expects the decisive return of the
orchestra at this point to close the exposition. Hepokoski and Darcy agree with this reading,
noting that R2 usually completes the ‘larger’ exposition, finding it ‘a virtually invariable
norm’ that S1’s rotation is incomplete without the R2 tutti (2006, p. 548; also cf. pp. 442 and
444).11

One piece that stands significantly behind Moscheles’s and Bennett’s works and
provides some precedent for their procedures, however, is Mendelssohn’s Piano Concerto No.
1 in G minor, Op. 25 (1831). Here the cadential trills in the piano that generically signal the
arrival of the R2 tutti and with it the close of the exposition are extended out at length over a
modulating passage to V/i which serves, in lieu of a development section, as a lengthy
retransition back to the opening and, with it, a recapitulation of the exposition’s material. The
design here points up the harmonic instability of the entire exposition, in which the most
stable statement of secondary material is in the tritonally distant D<flat> major, and whose
later reversion to the expected B<flat> (weakened through inversion) already shows distinct
signs of disintegrating back to G minor before the cadential section arrives.

Despite its evident influence, Mendelssohn’s differs crucially in that there is no real
sense of an R2 (or R1) being present: concerto ritornello form is almost entirely abandoned,
the movement approaching a pure sonata structure (albeit one truncated through the
suppression of any extensive development section into an abridged – type 1 – sonata design).

Unlike the design of Moscheles’s and Bennett’s concertos, the solo exposition is not
surrounded by the generic marker of the tutti in R1–R2-like function. In one sense,
Mendelssohn’s concerto is more radical, yet the closer adherence to some semblance of
traditional concerto form in Moscheles and Bennett creates scope for greater ambiguity
between the functional relationship of sections.

A more extreme suggestion along these lines, however, is given in Stephan
Lindeman’s reading of Bennett’s F minor Concerto. For Lindeman (1999, p. 202), the tutti
statement of P-material labelled R2 here is in fact the recapitulation (corresponding to what
other scholars term R3 – the optional ritornello demarcating the division between
development and recapitulation); Bennett’s movement simply bypasses R2 and the S2-based
development entirely. Though this interpretation is in some ways problematic, the reading
is nonetheless intriguing and provides a stepping stone for some useful further reflection.

Mendelssohn aside, it is quite extraordinary to have no R2 in a concerto first
movement. In fact there is only one other example in all 116 concerto forms Lindeman
analyses that also lacks the R2, and this is quite a special case – Cramer’s Concerto No. 8 in
D minor, which truncates the entire first movement following a unitary exposition and leads
directly into the slow movement. In other words, not a single concerto utilising ritornello
form misses R2 and continues after this point should have been reached – apart from
Bennett’s Fourth. Curiously enough, Lindeman gives quite a different reading to
Moscheles’s *Concerto pathétique*, noting that that this tutti – labelled here simply R2 – is 
back in the tonic C minor without further commenting on how unusual this is.\(^{14}\) What makes 
this interpretation more incongruent is that Moscheles’s C minor Concerto also drastically 
excises all P material from its recapitulation, just as Bennett’s seems to.\(^{15}\) To be consistent 
with this reading of Moscheles’s concerto – even ignoring Bennett’s dedication of his work to 
Moscheles and the dialogic reading posited here – Bennett’s F minor Concerto surely must be 
read in the same manner. And, beyond the ordinal anomaly, is it not slightly perverse for a 
concerto to have the rare R3 when no R2, virtually ubiquitous in first-movement concerto 
forms, is present?

As explained earlier, a tutti statement of P material here, following dominant 
prolongation after a display episode, will invariably be heard as forming R2 even if 
subsequently one might wish to revise this interpretation. In Bennett’s case especially, the 
rhetoric and structure of bars 235–257 in its restatement of P material is so indubitably that of 
a culminatory R2, which melts into a transition to a softer, solo-dominated developmental 
episode (see Ex. 1).

The first form bars form an extended sentential phrase marked by subdominant 
cadential tendencies (i–V\(^7\)/iv–iv–ii\(^\text{6/5}\)–V–i) and rhetorically intensifying descending scales 
in the strings as a counterpoint to the melodic material as heard in its initial presentation. The 
next six bars (247–252) form a post-cadential winding-down typical in function of the central 
stage of an R2 passage. And as if to confirm this argument, the subsequent movement to 
D<flat> major (<flat>VI) initiated by the orchestra in preparation for the ensuing solo-
dominated section (S2, bar 258) matches the pattern and key scheme for that found for this 
section in all three of Bennett’s previous concertos (Fig. 4).
The same principle also corresponds to Moscheles’s work, where the second half of the R2 clearly functions in a transitional role to the D<flat> opening the development – the more extreme <flat>II. What in both composers’ examples thus works against reading this central orchestral statement as starting a recapitulation (as in Mendelssohn’s Op. 25) is the fact that it functions exactly like a culminating R2 and immediately modulates to a distant key, just as is found initiating a development/S2 section (and totally unlike the behaviour of a recapitulation).

Nonetheless, Lindeman’s parsing of this movement has the virtue of a reading that seeks to explain the ambiguities and formal complexities in this piece, rather than to ignore their presence altogether. It is undeniable that the form of Bennett’s movement is severely truncated and any sense of recapitulation seriously compromised. Just as I contended that Bennett is in dialogue with Moscheles’s concerto, so are both surely responding to Mendelssohn’s First Concerto (and perhaps also to Cramer’s Eighth Concerto; see Lindeman 2007, p. 136). By locating the recapitulation at bar 235, Lindeman is able to account for a reprise of primary and secondary themes, which would otherwise not be heard as having this function; in a sense, if there is no recapitulation here, there is none at all. The F minor Concerto is also exceptional for Bennett in that the S2 section uses S material, rather than P (as in Concertos Nos. 1 and 2) or new material (Concerto No. 3), which supports the recapitulation reading slightly more.

More accurately, then, one might say that bars 235ff. would almost certainly be heard by a listener as forming an exposition-closing R2, and yet in hindsight this passage could also be interpreted as belonging to some (still problematic) form of recapitulation. Bennett seems to be deliberately playing with expectations in this movement, hence the structure can be
interpreted in different ways. Music conforming to the texture and formal position of a second ritornello is followed by the usual slow, solo-dominated section that begins developments in concertos of the time. As the movement draws to an unexpectedly premature close, however, both of these past stages might be retrospectively understood as multi-tasking for the recapitulation of first and second themes. Within an apparently unevenly weighted, end-directed design, a higher formal symmetry may be observed. There are issues in this argument that need to be addressed later, most evidently the ‘non-resolving’ nature of the recapitulation. Where Lindeman is coming from, though, is a retrospective reading of the form that seeks a logical explanation of the truncated later stages of the movement – synchronically, abstractly, rather than responding to the diachronic experience of the music – and he hits on a genuinely thorny aspect of Bennett’s piece. At any rate, he has no problem per se with a sonata exposition ending back in the tonic minor, as his reading of Moscheles’s Seventh Concerto demonstrates.

The points above, then, seem to substantiate the validity of reading this R2 in Moscheles’s and Bennett’s concertos as forming the close of the exposition, despite the tonal anomaly. This belief would militate against another possible (and in this context quite reasonable) reading – that of R2 as forming a rondo-return of the P (Primary) theme in a ‘Type 4’ sonata-rondo design. In their characterisation of sonata-rondo forms, the exposition is considered by Hepokoski and Darcy as completed before the return of the Primary theme, which marks instead the onset of the development or second ‘rotation’ of materials. Thus conceived, an R2 is incompatible with the return of a type 4 sonata (though elsewhere Hepokoski and Darcy note, rather briefly, the conjunction of sonata rondo with concerto-ritornello form in the finales of Mozart’s piano concertos [2006, p. 417]). Rather, within the operating terms of Sonata Theory, a P-based R2 would be typically described as a ‘rotationally inert [wild]card that can be placed onto the sonata table at any number of later
occasions’ (ibid., p. 482; see also pp. 556–7). In other words, the Primary theme does not necessarily participate in an ordered cycle of thematic material, and as a consequence its reappearance need not imply the start of a new rotation (the development). (The R2 thus exists, in Hepokoski and Darcy’s own inimitable terminology, in a state of ‘rotational limbo’ [ibid., p. 519; see also p. 529].)

Examples of this practice are given by Hepokoski and Darcy, an important example again being Mozart’s K. 491, whose secondary key of E<flat> darkens to E<flat> minor (bars 220–241) with the reintroduction of the ‘fateful’ opening motive, described by these two in characteristically emotive language (‘the larger expressive aim of the movement is to suggest being caught within a destructive whirlpool of a fatalistic C minor. One is assailed throughout by recurring nightmare-visions of the thematic representation of that threat, namely the R1:\P motto or idée fixe, which sprouts up in various parts of the form’ [ibid., p. 541]). Where Moscheles differs is that his exposition actually returns to the original ‘fate-laden’ minor tonic itself, which goes far beyond Mozart’s C minor Concerto.16

There is also a precedent in the concerto genre for a tonally extraneous opening to R2, but in all these cases the disturbance is almost immediately normalised. The most obvious examples are found in Beethoven’s Triple Concerto Op. 56 (1804–5) and the Violin Concerto, Op. 61 (1806), in which a dominant preparation of the secondary key is subverted by an interrupted cadence to VI at the start of R2; in both of these, however, this tonal shock is brief and the music soon reverts back to the true secondary tonality.17 This procedure is possibly foreshadowed in Dussek’s F major Concerto Op. 27 (published 1794) and finds later development in Ferdinand Ries’s Concerto No. 3 in C minor (see Lindeman 1999, p. 37). It might well be derived in turn from the routine interpolation of a colourful harmonic digression (commonly to <flat>VI) into the later stages of the Display Episode, in which the music humorously seems to get stuck in the wrong key before resolving (often through a
German sixth) to the true dominant of S. In the works of these composers, the harmonic intrusion is simply moved to a later stage of the form, thus intensifying the novelty value of the digression. Thus, like Ries (another of Beethoven’s younger friends and sometime pupil), Moscheles could be said to be engaging in dialogue with this scheme, only taking it much further.

One further possible reading that might come to mind is that this i–III–i progression for an exposition is actually quite common in the concerto when taking place in the first ritornello – the modulating R1 paradigm. In the context of the historical move in the concerto from the ‘double’ to unitary exposition both works here might be thought of as offering a larger exposition modelled on just the ‘first’ exposition’s tonal design. Admittedly this reading proceeds more from a perceived similarity in pattern rather than any plausible deeper motivation or probable origin for this design, and it does not take into account the evident expressive quality of the two first movements. All the same, it highlights the complexity encountered when trying to fit the concerto into a sonata-based model and the difference between imposing a formal mould synchronically onto the music and attending to the functional interrelationship between sections.

Issues in Sonata Theory

There are two particular issues with these pieces I would like to explore in relation to their explanation through Sonata Theory: first, the implications the tonal course of the their expositions have for the EEC (Essential Expositional Closure) and EET (Essential Expositional Trajectory) theorised by Hepokoski and Darcy; and secondly the relation their formal structures, especially post-expositional, have to their five sonata types.

Tonal issues in a non-modulatory exposition
Though professing the importance of thematic characterisation and ordering, and the generic norms and ‘dialogue’ with listener expectations behind this, the Sonata Theory model of the sonata is nonetheless extremely tonally orientated, positing fundamental tonal processes that drive to structural cadences in a manner at least partly indebted to the Schenkerian legacy of post-war American music theory. But how does one go about finding an EEC in an exposition that ultimately closes in the same key as it started? Does the concept of EEC even have any meaning in this context, since the fundamental point of the exposition is to articulate and confirm a large-scale modulation from tonic to secondary (i.e. non-tonic) key?

Two related cases that might appear to raise similar complications here are provided by the so-called ‘three-key’ sonata expositions often encountered in the nineteenth century, and the more unusual monotonal sonata expositions that are occasionally found. Hepokoski and Darcy theorise many apparent examples of the three-key design by means of their ‘trimodular block’ formulation, whereby one of the non-tonic keys is understood as structurally more important, thus preserving the deeper-level dualism between tonic and secondary tonality. This interpretation is borne out in many of the more celebrated examples in Schubert and Brahms, where the second tonality presented does appear to have a subordinate, dependent status on either the first or third keys, usually as a mediant relation within some type of paired ‘dual-tonic’ complex. Yet other examples do complicate the neatness of the dualistic structure Hepokoski and Darcy attempt to preserve so strenuously (Mendelssohn’s Quintet in A major, Op. 18, and Dvořák’s Quartet No. 10 in E<flat> are two cases in point).

More unusual monotonal sonata expositions can be found in a handful of works – Chopin’s Piano Sonata No. 1 in C minor (1828; entire exposition in C minor) and Piano Trio in G minor, Op. 8 (1829, again tonic minor throughout), and later in Franck’s Piano Trio Op. 1 No. 1 (1840; F<sharp> minor to F<sharp> major). In concertos, I am aware of only one
example: Chopin’s First Piano Concerto (1830, E minor to E major), although here the recapitulation of S is actually in III. In all these works the unusual tonal scheme might conventionally be explained away by virtue of their composer’s youth and lack of training in proper sonata etiquette (though a revisionary account of this procedure might also be intriguing). Clearly youthful inexperience cannot be levelled at either Moscheles or Bennett in the concertos under discussion here, and both moreover actually make the expected tonal movement away from the tonic only to subvert it at a later stage. If the teenage Chopin and Franck are both naïvely ‘mistaken’, Moscheles and Bennett are knowingly deformational.

The EEC is defined as the first PAC in the secondary tonality following the MC that goes on to new material; various additional requirements arise that complicate this definition, such as the subsequent return of a pre-EEC S theme nullifying the earlier EEC effect and, in the double-exposition concerto, the same occurring if an R1:S theme subsequently appears. The concerto (type 5) form in fact already presents a particularly slippery case for Hepokoski and Darcy since, as the authors note, the S section of many of Mozart’s concertos (their central paradigm) is ‘multimodular’, i.e. it contains multiple themes, and in many (though not all) the EEC is deferred until the arrival of R2.

Accordingly in Moscheles’s and Bennett’s concertos we are left with the problem that an early PAC within S1:S would constitute an EEC in a key other than the final tonal destination of the exposition (and therefore surely a paradoxical, if not contradictory, situation); and an EEC deferred until the arrival of R2 (and therefore according with the final tonal goal) would present a non-modulatory Essential Expositional Trajectory, which is impossible in Sonata Theory. But to persevere in this task, broadening the search for an EEC (or EEC-effect) as widely as possible to encompass both possible keys, one would need to find the first PAC in either III or i following the MC that is not followed by earlier material.
As it happens, this gives us a cadence to the tonic F minor midway through the R2 tutti in the Bennett, and strictly speaking none at all in the Moscheles. In this case, perhaps surprisingly given the above, Sonata Theory is still able to come up with a reasonable explanation that can adopt even these two unruly concertative cases (or ‘sonata stepchildren’) into its extended sonata family. In both, the music articulates a ‘failed exposition’, the erstwhile successful movement from tonic minor to relative major breaking down and with it the Essential Expositional Trajectory. In Moscheles’s work, the EEC is permanently deferred; in Bennett’s, the EEC effect in R2 is the sign of tonal ‘failure’. (Even an earlier EEC in relative major within S1, had it existed, could be explained by this theory as being subsequently ‘overridden’ by later tonal events.) Sonata Theory is sufficiently flexible to accommodate these pieces, and would provide a ready springboard for a type of hermeneutic reading Hepokoski favours: the pieces suggest an attempt (by the soloist) to break free of the minor-key pathos of R1 – which is shown to be unsuccessful even within the exposition, since the orchestra quickly restores the opening theme and tonality. In both, the familiar cadential prevarication around V/III preparing the R2 tutti affirmation of the triumphant relative major turns within a few short bars to chill as the dominant of the tonic minor creeps back in and the listener realises something is going wrong with the expected course of events. And furthermore, Bennett’s work is in ‘dialogue’ with Moscheles’s, and the latter in turn with Mendelssohn’s G minor Concerto and the C minor precedents by Mozart and Beethoven, the two later concertos forming tonally more radical deformations of this minor-major-minor trajectory.

There are, however, certain reservations about this theory that critics have raised that require further explication and disentangling.

i) The use (or abuse) of history
First is the historical basis of some of the key premises underpinning this theory, such as the issue of intertextuality just alluded to. The notion of ‘dialogue’ is central to the tenets of Sonata Theory. It has been claimed by sceptics of the system that the networks of influence which certain readings appear to rely on are not historically and culturally nuanced enough, but rather a recent theoretical construct which would not have been available to the composers in question. The criticism is largely directed against specific instances of Sonata Theory in practice or apparent claims on its behalf as to the scope of its remit, arguably falling short of affecting the possibility of its underlying validity qua theory. To this extent much of this criticism would seem potentially to be resolved by a more nuanced use of its concepts and greater qualification of its claims to historical grounding, without necessitating an overhaul of the entire system.

In the current instance it seems safe to hypothesise a nexus of influence in which these two concertos participated. While no conclusive evidence has yet come to light, it seems highly probable that Bennett would have known Moscheles’s Seventh Concerto by the time he started writing his Fourth Concerto. Besides the dedication, and the fact that the two composer-pianists lived in the same city at this time and were on friendly terms with each other and performed regularly together, there is a great deal of more detailed circumstantial evidence to suggest that Bennett would have been exposed to the earlier work. Moscheles’s *Concerto pathétique* was performed several times in London in the three years preceding the composition of Bennett’s Fourth. The first movement was played on its own on 1 May 1835, while the concerto was performed in its entirety a year later on 11 May 1836. It is possible Bennett might have seen either one of these performances; a few days prior to the second, having just turned twenty, he had performed his own Third Concerto at a Philharmonic Society concert. He would have missed a later airing on 30 May 1837, being in Leipzig at the time, but Moscheles’s concerto was performed again on 2 April 1838 at a Philharmonic
Society concert, with the composer at the piano and Bennett’s old teacher Cipriani Potter conducting – by which time the score had just come out.25 Would it not seem likely, then, that Bennett knew the piece by this stage? He was currently living in London and active as pianist and composer on the performing circuit. Indeed there are records of Bennett performing a duet in public with Moscheles in May 1838 (Moscheles’s *Hommage à Handel*, Op. 92), a concert which included Bennett’s own earlier F minor concerto. On 18 June Bennett performed the unpublished F minor work again at a Philharmonic Society concert, with Moscheles conducting.26 Bennett’s second concerto in F minor, Op. 19, was started and the greater part probably written that July on a summer break in Grantchester. It is tempting to conjecture that recent exposure to Moscheles’s Op. 93, allied to the dissatisfaction Bennett evidently felt concerning the earlier F minor work which he had just performed twice alongside Moscheles, may have been motivating factors as the 22-year-old composer attempted to fashion a more satisfactory F minor Concerto. Given how unusual the i–III–i ‘failed exposition’ is in first-movement concerto forms, the coincidence of two such pieces from the same geographic location at the same moment in time surely puts the case beyond reasonable doubt. Even the title Moscheles affixed to his work finds an echo in the epithet Bennett originally gave his concerto in the autograph score – ‘Appassionata’ – suggesting a doubling from their common lineage through Beethovenian allusion, an F minor counterpart to Moscheles’s C minor *Concerto pathétique*.27 One stage further back, the close personal association of both composers with Mendelssohn is so well known as to obviate the need for substantial comment. In both cases, too, there is clear evidence that both were acquainted with and greatly impressed by the G minor Concerto. Moscheles had witnessed Mendelssohn’s own performance of it at a Philharmonic concert in 1832, describing it as a piece which ‘sparkles with genius’ in invention and form,28 and, as Lindeman (1999, p. 35) has noted, ‘each of the three concertos that Moscheles
composed in the wake of Mendelssohn’s G minor Concerto, op. 25 … reveals his younger friend’s influence’. The young Bennett had similarly been exposed to Mendelssohn’s work as a student in London, as his son relates (Bennett 1907, p. 25).

The wider, ‘second-level’ network of now-canonic works of Mozart and Beethoven is also far less problematic applied as a compositional context to Moscheles and Bennett than, say, to earlier composers such as John Field. Although Horton has persuasively argued that the model of Mozart was not really available to British-based composers earlier in the century, performances of Mozart’s concertos picked up in London after 1819: Lindeman claims that ‘it is safe to assume all serious British composers were very familiar with their design from about 1825’ (2007, p. 105).²⁹ Potter was one of the leading advocates for Mozart’s concertos, and Bennett’s son leaves extended testimony to the special importance of this composer (Bennett 1907, pp. 22–6). The two minor-key concertos were among the most performed of Mozart’s in London in the first half of the nineteenth century; Potter performed it in 1836, Bennett himself eight years later (Ellsworth 1991, pp. 303–5). Moscheles had idolised Mozart as a youth in Prague and certainly knew Beethoven’s music well; in 1837, the year after he had premièred his *Concerto pathétique*, he gave a performance of Beethoven’s C minor (*ibid*, p. 282). Schumann, in his review of the *Concerto Pathétique* (1839, p. 6), in fact comments explicitly on this synthetic, ‘dialogic’ quality of Moscheles’s work:

In Moscheles we have the rare example of a musician who, although already in earlier years and even now has unceasingly concerned himself with the study of the old masters, has also observed the path of new appearances and has drawn from their advances. As he now commands every influence with his own innate individuality, so from this particular mixture of old, new and his own, arises a work such as the latest Concerto, clear and precise in form, tending towards the Romantic in character, and nonetheless characteristically original.³⁰

Here the applicability of the dialogic model is surely unproblematic.

One related charge concerning a lack of adequate historicisation has also been raised against certain elements of Sonata Theory, however, which is less easily solved. If it is
essential not only to relate ourselves to the composer’s horizon of expectations but furthermore to those of the audience who constructed meaning out of the listener’s contract with generic norms, it would help if there were verification that these concertos happened to be heard as formally unusual at the time (see also Riley 2008, p. 595). And on this matter there is far less supporting evidence. Simply staying with Schumann, from what can be ascertained from recorded observations the case is hardly overwhelming. Concerning Moscheles’s work, Schumann (1839, p. 6) merely comments:

\[ \text{The difference in form to others and Moscheles’s own earlier concertos will be immediately apparent to everyone. The first movement advances rapidly, the tuttis are shorter than usual, the orchestra imposes itself everywhere.}\]

This is a suggestive characterisation, but it does not directly address the deformation seen there – and this is as precise as Schumann gets concerning the first movement. And more worryingly, in a review of Bennett’s radically deformed Fourth Concerto, Schumann even claims that Bennett’s outer movements ‘offer nothing new in their shape’ (1840, p. 39). Contemporary reviews of these works in the London press were even less detailed. Moscheles’s \textit{Concerto pathétique}, for instance, is variously described as ‘a charming composition’ and ‘a composition (in the second movement especially) displaying great originality of thought, with ingenious, even masterly treatment’. This is the most we get; a later performance in 1838 following its publication elicited the response that ‘as we have it lying on our table for review, we shall reserve opinion to a future, though we trust early, opportunity’. This rather half-hearted promise seems not to have been carried out, and a year later Moscheles’s new \textit{Concerto pastorale} (No. 8) instead provided a new talking point. The applicability of this theory and reading to the sense of ‘historical listening’ which would seem to need to underpin it remains unproven.
A second point that arises in part from unresolved issues above concerns the actual
importance of tonality as a central element in determining form. Although Hepokoski and
Darcy distinguish between ‘tonal’ and ‘rhetorical’ form, pride of place is nonetheless given to
the former in questions of expository and wider structural closure. The absence of any
mention of tonality as a determinate of form in contemporaneous accounts, however, would
lead one to conclude that historical listeners understood the concerto in terms other than those
which Sonata Theory conceives as ‘essential’. For one crucial difference in the concerto
compared with the sonata, quartet or symphony, is the far greater emphasis placed on an
apparently secondary parameter – texture – as a central form-defining element. And one
idiosyncratic feature of the concerto that is insufficiently appreciated by Sonata Theory and
has a bearing on much of the discussion in this article is the extreme saliency of textural
structural markers such as the R2 tutti. Correspondingly, a theory that underplays these more
immediate, perceptible aspects of the music in favour of more esoteric qualities of cadential
closure potentially undervalues crucial features not only of individual examples but of an
entire genre’s form. As witnessed by contemporary critical reactions to Moscheles’s and
Bennett’s works, an understanding of concerto form less based on Schenkerian tonal
principles and giving greater importance to more readily audible parameters such as theme,
texture and gestural characterisation might reveal why these two works were not seen at the
time in quite such subversive terms as we would now view them.

With regard to Moscheles’s work, for instance, building on the blurring of orchestral
and solo functions already suggested in the model of Mendelssohn G Minor Concerto and
perceived by Schumann in his review, for the final module of the display episode (bar 158) a
variant of the opening motive of the primary theme is played by the soloist in a triumphant
E<flat> major (ff energico) before leading the music back to the dominant of C minor. In
key, theme and rhetoric this passage could fulfil the function of a modulating R2 tutti, but it is
played by the soloist and thus in textural terms still belongs to the exposition.\textsuperscript{36} It might be argued that the soloist has here anticipated the orchestra’s function, which raises the question of how much importance should be placed on textural contrast and tonal opposition relative to motivic return. Moscheles gives the listener the expected elements found at the end of a conventional concerto exposition; it is just that they are misaligned. But this fact was evidently not perceived as noteworthy beyond the new textural interaction. Such considerations would also help explain why contemporary writers heard nothing notably unusual in the first movement of Bennett’s Fourth Concerto. More loosely, then, concerto form might simply have been understood at the time as a flexible interaction between the textural opposition of tutti and solo, a variety of thematic ideas, and (usually) some degree of tonal contrast; if expected elements from all three are present to some degree, even if their exact interaction might be unusual, the individual concerto was normally accepted as unproblematic.

Applied to the concerto, then, Sonata Theory would seem to need greater flexibility in dealing with certain parameters, as well as a loosening of its dependence on tonal processes. And even if we accept the grounding tonal premises of Sonata Theory, the notions of EEC and EET highlighted by these two pieces call for further scrutiny. As mentioned, Hepokoski and Darcy note that the position of the EEC can be extremely varied within the first movements of Mozart’s piano concertos, occurring sometimes immediately after the first second-group theme, possibly midway through the solo exposition, or even after its end in R2. Yet it is questionable whether this variable placing really reflects major differences in these pieces, beyond the mere fact of having a PAC in one piece where an IAC exists in another, or a pre-PAC theme returning at a later stage (as happens with Bennett’s Fourth Concerto).

To give an example, in Moscheles’s previous six concertos a PAC occurs after the initial S theme, before a long string of ‘secondary theme’ modules that comprise what is
generally named a Display Episode (hereafter DE) leads in a dynamic curve of intensity to
R2, heard without exception as the culmination of the exposition. By this rule, the DE should
be termed ‘Closing-theme zone’ material, occurring after a satisfactory PAC in the secondary
tonality. Yet in a work (such as the present *Concerto pathétique*) without a PAC in this place,
or one in which the DE uses material sufficiently akin to S to cast doubt on the erstwhile
intervening EEC, the entire DE must necessarily be termed as belonging to S, and R2 on its
own as forming a closing zone. The Third Concerto in G minor, Op. 58 (1820), is already a
debatable case of the latter, in that the post-cadential passage immediately following the EEC
at bar 227 uses material derived at least in part from the second phrase of S (bars 71/191), and
leads directly into the DE without any further cadential break. Although I feel (partly on the
basis of Moscheles’s practice elsewhere) that one can draw a line at bar 227, it is a decidedly
slippery matter whether the theme here is sufficiently different as to constitute new material,
and thus to confirm the EEC it initially seems to articulate. If not, then since this post-
cadential theme is reused, slightly altered, as the R2 theme, expositional space might in turn
be thought of as being kept open not just until the PAC at the beginning of R2 at bar 303 but
until the PAC in B<flat> towards its end at bar 321. This reading, however, is quite at
variance with the rest of Moscheles’s practice, and does not to my mind reflect genuine or
salient differences in this concerto compared with his others.37

Bennett’s practice here is particularly a case in point. All three of his earlier concertos
provide subtly different cases illustrating the interaction between cadential closure and
rhetorical form. In his First Concerto the EEC may be placed at the opening of S²/DE (F
major, bar 131), as with Moscheles’s examples, but in the Second Concerto the EEC occurs
with the opening of R2 (B<flat> major, bar 275), owing to the return of S1 material in bars
240ff. following the DE. In the Third Concerto it reverts again to the start of S²/DE (E<flat>
major, bar 181), but then material from the S¹ theme returns in R2 (in conjunction with P¹).
Does this return of S material in R2 in the Third Concerto retrospectively invalidate the earlier structural cadence? Given the slippery identity of R2, and especially the fact that (as Horton notes) it can be split across the higher-level sonata functions of exposition and development, how do we know when the exposition ends? With the capacity for retrospective reinterpretation of former cadences, any reappearance of S material in an S tonality might reopen expositional space – even (theoretically!) if this should occur in what might otherwise appear to be the coda of a sonata form – unless some other definition of expositional closure is called upon. To define the expositional close in terms of a structural cadence that can vary its final position according to where the exposition may be said to end gets us caught in circular reasoning. In practice, such confusion is seldom the case; but potentially it may happen, and the concerto’s ritornello-sonata fusion is a clear locus for such theoretical dilemmas.

It is also arguable whether the difference between the coincidence of an IAC and PAC is a hard-and-fast rule worthy of causing such a massive formal and terminological distinction. Often, indeed, it is difficult to hear whether a V-I cadence has ≥1 or ≥3 or ≥5 in the uppermost voice, especially in orchestral music. So much might depend on whether the composer chose to double a particular line at the octave in the flute, thus forming the highest voice, or whether, despite a cadence to ≥1 in what is technically the uppermost voice, scale degrees ≥3 or ≥5 are far more emphatically heard due to deliberate choices in scoring and doubling. This is an issue exacerbated in the concerto genre, where two quite distinct textural streams are given. The solo part may, say, contain an IAC but the orchestra a PAC. A pertinent case occurs at the start of the Coda/R4 of Moscheles’s *Concerto pathétique*: the orchestra completes a PAC in C minor in bar 319-20, with strong root motion g–c in the lowest voice formed by the cellos. Yet the soloist reaches the tonic chord through contrary motion scales, touching the C
an octave lower from its upper neighbour D, thus technically making the entire progression an
IAC (see Ex. 2). I would claim, however, that there is a sense of completion here – certainly
in comparison with the numerous IACs encountered earlier in Moscheles’s movement –
which justifies describing the cadence as a PAC (and therefore as forming the ESC).

[INSERT EX. 2 HERE]

The demand for an EEC or ESC may sit rather uncomfortably alongside the historical
and ‘listener-orientated’ models of form and expectation Hepokoski and Darcy frequently call
upon. The general principle, taken from William Rothstein (1989, p. 116), that closing theme
status should be conferred on material occurring after ‘the first strongly articulated perfect
cadence in the goal key’ (and which is not distinctly followed by earlier S material), seems
reasonable as a general rule of thumb, but in practice other parameters such as register and
timbre play so great a role in how forcefully this cadence is expressed as to beg for their
inclusion in any sensitive account of an individual work. The EEC often has only a score-
based existence, sometimes at odds with the aural experience. Unless the listener has perfect
pitch, and then is able retrospectively to defer or relocate this all-important cadence, the EEC
will remain an abstract entity.

In itself this is not so serious a charge – after all there is no reason why an analytical
construct is false just because it cannot be heard – but it seems strange that the formal
function of such a generic feature as a Display Episode should change from secondary-theme
to closing-theme status just because of a single note in a preceding cadence. Given the
ubiquity of the secondary-theme format in nineteenth-century concerto movements – that is, a
lyrical second theme followed by multimodular DE culminating in R2 – the variable
positioning of the EEC seems to be of slight importance. An obsession with tonal trajectory
and closure smacks of Schenkerian metaphysics that is nowadays far less readily accepted as
a matter of faith. After all, the more immediate, audible function of the multimodular
display episode is not to prolong an underlying harmonic tension so much as to offer an
opportunity for virtuosity; the historical listener (and composer) may have been more
interested in moment-to-moment events than in any tonal goal over the span of modules.

Essentially my argument is putting forward a case for a ‘soft’ version of Sonata
Theory that is both flexible to the nuances of the individual analytical case and aware that
tonality may well not be the most significant means of formal articulation in genres such as
the concerto. It would be unjust to criticise Sonata Theory too harshly here, since it was not
developed with nineteenth-century concerto form uppermost in mind. In limiting their
discussion of the Type 5 Sonata to a provisional exposition of the Mozartian model,
Hepokoski and Darcy recognise that the concerto is a tricky and apparently hybrid genre; it is
therefore not altogether surprising that a Sonata Theory model might require modification in
treating nineteenth-century examples.

Multiple sonata-type deformations

Moscheles’s C minor Concerto, as said, is not just a possible merging of sonata Types
5 and 3, but also contains elements of Type 2 (binary) and even Type 4 (rondo). Turning our
attention to the second half of the movement also suggests the possibility that the R2 tonic
statement of the Primary theme might in some form substitute for or balance the excised first
group recapitulation. In Hepokoski and Darcy’s terminology, the development space is
‘pasted over’ two possible halves of the recapitulatory rotation. Thus this almost approaches
a 2-part exposition-recapitulation, with a developmental interpolation in the recapitulation –
i.e., a Type 1 sonata.

The idea that one can just ‘insert’ extended passages like this might be objectionable
to some analysts, but Sonata Theory provides rationale for this procedure in its focus on
thematic ordering – sections as large-scale ‘rotations’ of a model grouping provided by the
exposition. While hearing R2 as starting a recapitulatory rotation goes against probability
given generic expectations, and the idea of retrospective reinterpretation might appear to be
just that type of analytical opportunism decried earlier, this reading both touches on the
balancing effect of this post-MC tonic P return in relation to the curtailed binary
recapitulation, and points forward to Bennett’s procedure in his F minor Concerto. There, the
post-expositional stages of the movement are even more drastically broken down, the R2
return of P in the tonic leading without any real break to a culminatory coda. While there is
no clear subsequent recapitulation point (unlike the Moscheles), the ensuing extended
statement of S in D<flat> major forms some kind of recapitulatory symmetry to the
exposition’s A<flat> major statement (aligning it with Lindeman’s reading noted earlier).
What further adds to the plausibility of this interpretation is the structure of Bennett’s
subsequent (and final) concerted work for piano and orchestra, the Sixth Concerto in A minor
(1843), originally entitled Concert-Stück, whose first movement in Peter Horton’s description
omits the recapitulation altogether and moves ‘directly from the development into a short
coda’ (2007, p. 140).43
This idea of ‘balancing mediant/submediant’ is admittedly a thorny issue; Hepokoski
has argued that rather than the submediant substituting for the tonic, the non-tonic reprise of
secondary material in such works may be heard instead as enacting recapitulatory ‘failure’
(see Hepokoski 2002a, 2002b, and Hepokoski and Darcy 2006, pp. 242–5). A direct parallel,
featuring the same keys (P in F minor, S in A<flat> and recapitulated in D<flat>), can be
found in Beethoven’s Overture to Egmont, another work which Bennett is potentially in
dialogue with in his F minor Concerto. Does Bennett’s piece, then, contain a failed
exposition and failed recapitulation? That would presumably make it a failed sonata,
simplyciter. As Hepokoski and Darcy note in relation to the (rare) failed exposition
deformation, ‘in all cases the subsequent recapitulation is deeply problematised’ (2006, p. 179). In fact this concerto presents an intriguing case. Drawing inspiration from an earlier reading of Hepokoski (2007), one might say that there exists a ‘greater’ and ‘lesser’, or ‘major’ and ‘minor’, version of the second subject’s theme. The former is the true secondary theme heard in A<flat> in the exposition (S, bars 132–166), but this is actually prefigured in the opening orchestral tutti in a more nascent, minor-hued form (the ‘lesser’ variant, P, bars 46–50), a motivic idea rather than fully fledged theme (Ex. 3).  

[INSERT EX. 3 HERE]  

According to generic expectation and the manner in which the idea is presented, the listener would expect a triumphant minor-to-major trajectory, a development from imperfect potential to perfect actualisation, but in the aftermath of the R2 tonic minor complications the major version is never heard in the tonic F. The upshot is that the ‘lesser’ minor version from the opening orchestral ritornello reasserts itself in the extended passage from bar 308 and takes over the movement; the association of ‘strong’ and ‘weak’ versions have been turned on their head, with disastrous consequences.  

Thus although S-related material has been heard in the tonic minor in the post-expositional stages (the PS theme), this was merely an idea that had originally been heard in this key, while the true S theme that appeared in a non-tonic key has not been resolved. As much as there is a recapitulation in this piece, it would appear to be a failure.

At a larger level, then, the expected development section takes on recapitulatory properties, and in addition to the R2 restatement of P, could thus be analysed as a developmental-recapitulation rotation, resulting in a Type 1/2 mix sometimes termed a ‘Bruckner deformation’. Hence, on paper, Bennett’s work can be read as embodying
characteristics of types 1, 2, 4 and 5. In fact, the type it is least similar to is type 3 – the
normative type concerto first movements were supposed to be approaching at this time
according to the historical narrative set out at the opening of this article.

One final criticism that could be directed against this analysis is: why do we need to
run through an inventory of all five sonata types, showing how it is a bit like one but in other
respects incompatible and not completely like any? For all the time and effort spent trying to
define the form against Sonata Theory’s five types, is the unusual yet logical and compelling
structure of these two movements not just as well expressed in the simple diagram given
earlier? This might be true, though Hepokoski and Darcy might counter that it is only by
comparing a form with their 5 ideal types – or some ideal type – that one can justify calling it
a sonata deformation and not simply a non-sonata structure.

To conclude in suitably meta-theoretical vein, one might make some brief remarks about the
applicability of Sonata Theory to the type of concerto form witnessed in these two concertos.
Moscheles’s Seventh and Bennett’s Fourth Concertos raised problematic issues; flexibly used,
however, this theory seems able not only to cope with the challenge thrown by these two
works but moreover to come up with a musically and expressively sensitive interpretation of
both drawing on this divergence from theoretical norms. Indeed, the particular tonal
difficulty raised by these pieces is inevitably going to be exacerbated within Sonata Theory;
the tonality of R2 could only become such a major problem for a sonata-based concerto first-
movement model, especially one with the need for a structural cadence located after S in the
EEC’s manner, and that places R2 in the exposition, encountering a work with no EEC in S1.
One might thus be tempted to look elsewhere for theoretical formulations; if a ritornello-
based model were used for the two concertos, for instance, this issue would not arise so
drastically. Conversely, though, the sonata-theory reading brings to light just this feature
which is surely the most salient aspect of the movement, structurally and expressively. As
Hepokoski states, Sonata Theory’s method ‘can be placed into relief by examining extreme cases’ (Caplin et al, 2009, p. 73). One might say that ultimately the criterion for determining the success of a theory is in the insights it can throw on individual pieces; the proof of the theoretical pudding is in the analytical eating. And here Sonata Theory certainly comes up with something. Yet despite this, there remain certain issues with the deeper theoretical underpinning and its application in analytical practice, especially concerning the relation of historical and presentist claims and the importance of tonal closure, which will probably continue to be disputed.

With regard to the concern highlighted at the start of this article concerning the historiography of the piano concerto, the overall argument supports the implications of the work undertaken by Lindeman and Julian Horton, that the history of the concerto outside the Mozart-Beethoven line may be more varied and prone to innovative formal procedure than often assumed. It points to the importance of knowing this supposedly peripheral repertoire in coming to an adequate understanding of form in the nineteenth century. At the same time, in terms of specific networks of influence we see how these two works may provide a via media between the Viennese and London models of concerto form. To understand Bennett’s concertos and their dialogic interlocutors one must know Moscheles as much as Mozart, Cramer as well as Mendelssohn and Beethoven.

And at the broadest level, what becomes especially intriguing from studying the compositions of such figures as Moscheles and Bennett is how many of those works which ostensibly would appear to move to a simpler, more purely sonata-like model actually result in more complex designs – i.e. the correlation between single-exposition concerto form and extreme departures from norms. Often such pieces are combined with other traits such as the running-on of movements, multi-functional (3- or 4-in-1) designs, truncated first-movement deformations (such as the first-movement exposition leading directly to the second
movement) and overt thematic links between movements. In other words, some of the most intriguing structural experiments in music from the early nineteenth century happen in the concerto and *Konzertstück* genres, as they overlap with sonata principles. In turn, such formal conceits increasingly find their way into the sonata structures of symphonic and chamber music in the second half of the century. Much scholarly effort has gone in to showing how first-movement concerto form was historically not derived from the sonata (as the later nineteenth century saw it) but independent, possibly earlier in origin, at best based on shared principles about tonal behaviour, resolution and structure. This very independence would help explain why the genre could, conversely, have provided a means for innovation in the sonata, as being something separate, a new fertilising element. It has often been disputed precisely how and when the sonata infiltrated the concerto. But it might be even more intriguing to argue that the influence could also have occurred the other way round.
NOTES

The groundwork for this paper was carried out as part of a project on sonata form in the early nineteenth century funded by the Irish Research Council for the Humanities and Social Sciences and University College Dublin. I would like to thank Julian Horton and Paul Wingfield for their support, Susan Wollenberg for her comments on a later draft of the article, and the anonymous readers for this journal.

1 See especially the papers devoted to the question of *Formenlehre* at the Sixth European Music Analysis Conference, Freiburg, 10–14 October 2007, and the resulting three-part exchange between William Caplin, James Hepokoski and James Webster (2009).

2 Bennett’s concerto should not be confused with an earlier F minor Piano Concerto written in 1836 (his fourth composed, but unpublished). Both this work and the one discussed here in fact share the same movement – a Barcarolle – which, even more peculiarly, was introduced into each work as a replacement for its original second movement. See Bennett (1907), pp. 42 and 75–6, and Williamson (1996), p. 361.

3 One might note in this context, though, that the ‘virtuoso problem’ afflicting the critical standing of other pianist-composers in subsequent reception is not an acute issue with these two. (Robert Schumann, one of the primary sources for later value-judgements against the music of virtuoso pianists, had great respect for both.) Unlike such figures as Daniel Steibelt, Henri Herz or Friedrich Kalkbrenner, Moscheles and Bennett do not need to be saved from the charge of meretriciousness – merely from a dusty oblivion.

4 To their credit Hepokoski and Darcy never imply this reading: rather, they suggest that formal models coalesced around ideal types constructed (possibly *ex post facto*) from newly canonic works by Haydn, Mozart and Beethoven; by contrast, the instrumental structures of later eighteenth-century *Kleinmeister* are possibly more varied and lax (Hepokoski and Darcy
Following Mark Evan Bonds (1991), structural norms are not a liability but rather a necessary condition for the ability to construct musical form in the first place.

5 The R2 tutti, as common with Moscheles, is tonally mobile, starting in C minor but ending on V/D<flat>. This modulation could, in theory, confuse the question of which key the exposition ends in. In practice, however, there is no ambiguity either here or elsewhere in Moscheles’s works. In all six of Moscheles’s previous concerti the PAC that serves as the EEC occurs some time earlier within the S1 exposition, at the beginning of the so-called Display Episode, and the cadence completed with the opening of the R2 tutti is in every case into this same secondary tonality as the exposition. Here in the Concerto Pathétique there is no suggestion anywhere in R2 of the E<flat> secondary key, the tutti’s opening completing an IAC into C minor before moving further afield to V/D<flat> in preparation for the key initiating the S2 / Development section (a procedure followed in all of Moscheles’s previous examples save for the Concerto No. 3 in G minor, whose tutti remains in the secondary totality of B<flat>). Neither has there been at any earlier stage in S1 a PAC into the secondary key of E<flat> that could serve as a satisfactory EEC.

6 See further Williamson (1996), p. 92, which reproduces a letter from Bennett to Moscheles requesting permission for the dedication (5 June 1839). I acknowledge that dedication need not imply musical influence; given the context of their peculiar shared structure, however, Bennett’s honouring of Moscheles on the title page of his concerto adds to the plausibility of the connection as being conscious.

7 Bennett studied at the Academy from 1826 to 1836. Moscheles’s concertos were in fact twice as frequently performed as Mozart’s in London at this time: Therese Ellsworth lists 22 recorded performances of Mozart Piano Concertos for the period 1821 to 1840, compared with 44 for Moscheles (Ellsworth [1991], Appendix 2, pp. 301–5).
The examples given are from Czerny (*School of Practical Composition*, 1848) and (admittedly much later) Ebenezer Prout, who, following Czerny’s lead in his *Grove* dictionary entry, sees R2 as leading to the development and later (in his *Applied Forms*) more explicitly as commencing this section.

Cf. Hepokoski and Darcy (2006), p. 559, on this type of linking R2.

In the Concerto No. 3, Op. 9, R2 synthesises motives from both Primary and Secondary groups before continuing with Primary material.

Also compare with the model derived by Rink (1997), pp. 4–5, from Amster (1931), cited by Horton (2011), pp. 53–4, which gives R2 as belonging to the exposition. Claudia Macdonald likewise interprets the function of the second tutti in the virtuoso concerto as being ‘to signal the end of the exposition and lead to the key of the second solo’; see Macdonald (2005), p. 20.

The table is also reproduced in Lindeman (2007), pp. 132–3. The structure of Bennett’s movement has caused some confusion for its few previous commentators. A similar interpretation has been given by Peter Horton (2007), p. 135, who speaks of ‘no real development section’ and the central orchestral tutti being omitted. Nicholas Temperley meanwhile seems to refer to the D<flat> statement of S material from bar 258 as being a recapitulation, but nevertheless mentions a development section in the movement (Temperley [1959], vol. II, p. 200). Hans Engel (1927), p. 237, analyses the movement as I have, speaking of the exposition’s second theme appearing in A<flat> and then F minor, followed by a short R2 tutti ‘rather ineffectively in the tonic’ (das hier nicht sehr wirkungsvoll in der Tonika steht) and extremely ‘untended’ (vernachlässigt) development consisting of the second theme in E<flat> [recte: D<flat>]. Engel, for reasons best known to himself, considered ‘Bennet’ [sic] an honorary German.
As noted, Mendelssohn’s Concerto No. 1 has no R2 but further lacks any distinction between solo and ritornello sections in its exposition. Here Lindeman reads the harmonically open-ended exposition as eliding with a brief transitional development (1999, p. 89).

Ibid., p. 250. Hans Engel again accords with this reading and mine (1927, p. 208). A far milder precedent for the tonic return in the middle of the first-movement’s structure may be seen in Moscheles’s preceding *Concerto fantastique*, which closes unproblematically in the mediant major (a departure from his earlier practice, though quite consistent with Beethovenian sonata norms) before modulating to the tonic minor in the second half of R2 in preparation for the start of the development/S2 rotation (P:B<flat>; S:III; R2 III→i). A source for this procedure may well have been Kalkbrenner’s popular Concerto No. 1 in D minor, Op. 61 (1823), whose R2 tutti, after confirming with its first chord the secondary key of F major promised by the preceding cadential trills, swerves immediately back to V/i (bar 209). While prophetic of Moscheles Seventh and Bennett’s Fourth, the initial PAC in III completed by the orchestra, allied with the EEC given at the start of the Display Episode in bar 179, still preserves the tonal articulation expected of an exposition.

All that is heard is four bars of the S1 Exordium from the solo entrance (bars 254–257, originally at bars 56–61). Since no S material appears in R1 and no P material in S1 (i.e., Moscheles writes a unitary exposition) it is debatable whether this brief passage – a new idea, leading on directly to a transition – constitutes primary material.

A variant of Mozart’s procedure is found in Beethoven’s own C minor Concerto, where the relative E<flat>, quickly reached in the opening ritornello, clouds over to the minor (bars 37–47). E<flat> minor is also heard as a brief modal contrast within the later stages of S1 (the post-cadential closing section), though here there is no motivic connection to the opening C minor theme.
The Triple Concerto is more extreme in that the secondary tonality is the rather unusual submediant, A major; and though R2 reverts back to this tonality for the start of the development, it is now in the opposing minor mode.

The fact that there are only two, early examples of a modulating R1 in Mozart (K. 413 and K. 449) has long skewed the interpretation of the concerto form; in addition to the better-known examples in Beethoven’s first three piano concertos, this practice was widespread. Dussek seems to have been one of the earliest prominent proponents of this design; further examples can be found in Cramer, Field, Steibelt, Hummel and Moscheles. See Lindeman (1999), p. 24, and Horton (2011), pp. 65–70, esp. table 13 on p. 70.

Claudia Macdonald (2005, pp. 42–3) discusses an early sketch of an unfinished F major Concerto by Schumann from 1830 whose exposition similarly returns to the tonic. In a wider context, Charles Rosen has also pointed to Boccherini’s A minor Quintet Op. 25 No. 6 (1778), which moves to the relative C but then back to the tonic major for the close of the exposition, a procedure deemed ‘a singularly lax misunderstanding’ by Rosen (1988, p. 245).

Moscheles’s movement is notable for how its numerous possibilities for expositional closure are overridden. In a departure from his practice in all six earlier concerti, the texture and harmonic template suggestive of the generic Display Episode beginning (in the wrong key, A major) around bar 117 is not set off from the preceding music with an EEC-forming PAC but continues from the jaunty dotted idea of S² (bar 100). Ensuing new Secondary-theme modules at bars 145 (S³, derived from an idea introduced in passing at bar 97) and 158 (S⁴) are similarly underarticulated through IACs; each in turn sounds like a further prevarication of expositional closure. The latter in fact might have been suited to providing material for an R2 tutti: although repeating the opening harmonic template of S³ (I–vi–o⁷) and thus audibly forming a continuation of the Secondary group, the motivic content is closely related to the Primary material. However, as with the start of S³, the brief suggestion of an
orchestral entry at this point is quickly quashed as the soloist continues, and the music darkens, slipping back to V/C minor. Nowhere is a PAC actually given.

21 See the discussion in Hepokoski and Darcy (2006), pp. 177–9 & 190–1; the i–III–i tonal scheme is not given by the authors in any of their examples but the procedure of Moscheles’s and Bennett’s works certainly fits into this category. The closest predecessor is the first movement of Haydn’s Op. 20 No. 3, which after a lengthy S area in the relative major peters out without having reached any EEC onto chord vii of the tonic G minor in preparation for the return of that key. Probably more relevant as a model (if arguably less formally analogous or striking than the Haydn example) is the first movement of Beethoven’s F minor ‘Appassionata’ Sonata (see ibid., p. 312). Although its exposition does modulate to the relative major, A<flat>, this soon collapses back into its parallel mode, A<flat> minor, that had been hinted in the transition (bars 25–34). This may clearly be read as a ‘failed’ exposition in expressive terms, even though it does modulate away from the tonic and is provided with a (weak) EEC (bar 61). Moscheles and Bennett greatly increase the deformational stakes in their return to the tonic minor.


24 On 25 April 1836, The Musical World, 1/7 (29 April 1836), p. 110. Bennett was away for the second half of May attending the Lower Rhine Festival in Düsseldorf – his first of several visits to Germany.
The Musical World, 5/64 (2 June 1837), p. 186, & 8/14 (5 April 1838), p. 232. The piano score was advertised as being ‘in press’ in March 1838 (Musical World, 8/10 [8 March 1838], p. 168) and was in hand for the reviewer of the April concert.


See Williamson (1996), p. 84; Williamson cites Martin Outhwaite (1990) as having earlier noticed this parallel. As noted, both concertos might be seen as being in dialogue with the harmonic design of the exposition of Beethoven’s Op. 57 (a milder i–III–iii EET breakdown).

Moscheles, diary entry for 28 May 1832, in Moscheles (1873), vol. I, p. 271.


‘Das Abweichende in der Form mit andern und Moscheles’ eigenen früheren Concerten wird Jedem im Augenblick auffallen. Der erste Satz schreitet rasch vorwärts, die Tutti sind kürzer als gewöhnlich, das Orchester greift überall mit ein; […]’

‘Die andern Sätze [the outer two movements] bieten nichts Neues in ihrer Gestaltung, oder besser gesagt, sie suchen das Neue nicht im Auffallenden, sondern eher im Anspruchlosen’.

The example here – drawing on just one member of Moscheles’s and Bennett’s audience (albeit an extremely eminent one) – is obviously extremely limited; a far wider survey would be needed to make any proper claim as to the efficacy of Sonata Theory. But it is revealing
that such perceptions of Bennett’s work from such a sensitive musician – and moreover one so favourably disposed towards the composer – were possible. As Nicholas Temperley (1989) has argued, Schumann undoubtedly had a very high opinion of Sterndale Bennett.


34 As the examples above imply, the capacity for comprehending music with a 20th-century type of analytical acumen (in which such ideas of dialogic choice may become apparent) might not have been available for 19th-century audiences; the methodological premises of Sonata Theory could be said to be foreign to those of 18th- and 19th-century musical perception. See Riley’s shrewd line of attack (Riley 2008, pp. 593–8).


36 As argued above, the fact that the cadence at the start of this theme is merely an IAC nevertheless undercuts the audible sense of attainment; the passage from bar 158 could easily have been used, slightly modified, as an R2, but this lack of cadential definition, along with its close similarity to the S3 module immediately preceding, still ties it in with the solo exposition.

37 I acknowledge that Hepokoski and Darcy are sensitive to the need for individual discernment in such questions (see Hepokoski and Darcy [2006], pp. 151ff). Indeed, one of the strengths of Sonata Theory is that it recognises these very aspects of thematic and harmonic design that in many cases can shed useful light on the music’s characteristic qualities. I think of the exposition of Beethoven’s Sonata in G major, Op. 31 No. 1, where
Sonata Theory highlights just this equivocation in the Secondary group between modules that are a sufficiently close variation of the preceding material to constitute a continuation of S and those that have become distant enough to count as belonging to a Closing zone. There is no clear demarcation, and this is one of the noteworthy aspects of Beethoven’s exposition (as Dahlhaus, or Schmalfeldt, might say, secondary material becomes closing material).

However, as argued above, the concerto genre raises such problems far more regularly.

38 This is not just a speculative objection. I needed to consult the orchestral score to be totally certain that the cadence at bar 158 of Moscheles’s Seventh was in fact an IAC, not a PAC (at 5'06" in the Howard Shelley recording (Hyperion CDA67385: The Romantic Piano Concerto vol. 32); the piano score gives in the piano’s upper voice, which does indeed turn out to be the highest pitch in the orchestra (flute 1), though and in the violins come across more strongly).

39 As Caplin has pointed out, the idea that a sonata exposition must contain a single ‘essential expositional close’ ‘is as dogmatic an assertion as to be found in music theory from any era’. Caplin, ‘Response to the Comments’, in Caplin, Hepokoski and Webster (2009), p. 53; see further the criticisms offered by Caplin and Webster on this point (ibid., pp. 29–30, 137–8).

40 In fairness to Hepokoski and Darcy, some provision is given for flexibility in such EEC / Closing Theme issues in extreme cases where musical rhetoric may override their posited structural cadences (Hepokoski and Darcy [2006], p. 191).

41 Hepokoski (2012), pp. 217–51, has recently provided an account of Brahms’s First Concerto drawing on the terms of the type 5 sonata set out in Elements. However, as he notes, Brahms is calling upon earlier historical models in providing a double exposition, and thus this piece is not especially representative of the nineteenth-century type of concerto form considered here.
The underlying design hence reveals its clear kinship with Mendelssohn’s Op. 25 Concerto. Also note Claudia Macdonald’s analysis of the movement (Macdonald [2005], pp. 157–8). Her parsing of the exposition (based only on harmonic function) clearly differs from my reading but her interpretation of the second half of the movement certainly accords with the point made above.

The concerto was rediscovered only in 1992 (see Williamson [1994], pp. 115–29); I have not been able to study this work personally, the manuscript being in a private collection. Peter Horton reads this design as responding to Moscheles’s previous Concerto, No. 6 (the Fantastique, Op. 90 [1833]), which implies an intriguing network of influence between these two composers especially since several commentators have suggested that Bennett’s initial four-movement conception for his First Concerto may have influenced this concerto of Moscheles dating from the following year.

In terms of pitch structure the two ideas have little in common, with the exception of the leaping dotted-note tag at the end of each, but through this characteristic feature and the exact replication of rhythm the two are clearly related. Still, the lack of diastematic commonality justifies the designation of the first, minor version as belonging properly to the P group of themes, with which it shares several important traits (the rising arpeggio, the dotted figure in more embryonic form), hence the designation PS. This labelling also implicitly disagrees with double-exposition readings of the movement, which are only justified by appeal to this brief F minor idea as being already properly ‘second subject’. It is worth noting that in Bennett’s three earlier concertos (which indisputably do feature a double exposition) the statement of the S theme in R1 is in all cases in the secondary tonality – the modulating R1 normative for English composers – whereas in Op. 19 the passage in question, thematically as close to P as it is similar to S, is in the tonic minor.
One might even say that the music’s very dwelling on the limpid beauty of this major theme, its reluctance to move on and let this enchanting idea go, is the cause of the movement’s downfall. A PAC at bar 166 following a quite extensive double statement of the secondary theme seems to suggest a safely attained EEC, but then, after a Display Episode (bars 174ff) with the music still not in any hurry to reach the safe pillar of R2’s tutti, the theme re-emerges above the piano arpeggios in the flute (bar 193\textsuperscript{4}), thus ‘reopening S-space’ and negating the previous EEC effect. The cadential prevarication continues, the coy avoidance of the expositional A<flat> major goal drawn out to yet greater lengths, until with the ominous entry of V / F minor everything darkens and collapses the music back into the tonic minor realm; it is now too late.

A list of works up to the 1840s featuring the intersection of these ‘progressive’ traits could include Spohr’s Violin Concerto No. 8, Cramer’s Piano Concerto No. 8, Weber’s *Konzertstück*, Mendelssohn’s Piano Concertos Nos. 1 & 2 and Violin Concerto, Alkan’s two *Concerti da Camera*, Clara Wieck’s Concerto in A minor, Herz’s Concertos Nos. 2 & 4, Moscheles’s Concertos Nos. 6 & 7, Bennett’s Concerto No. 4, and Schumann’s Piano Concerto in A minor. The overlap with the fantasia genre is also highly relevant to this discussion. It should be noted here that beyond a brief suggestion of the opening repeated march figure from the first movement at the start of the second in Moscheles’s work, neither of the pieces discussed in this study have notable thematic connections between their *opening* movement and the rest of the piece (unlike Moscheles’s previous concerto), although the latter two movements of Moscheles work are closely connected, the A<flat> major first theme of the second movement being reused in C minor form as the principal idea of the finale.
Abstract

While neither the music of Ignaz Moscheles nor that of William Sterndale Bennett feature in
the modern repertory, the piano concertos of both composers form part of a significant body
of works from the early nineteenth century that offer structural models at variance with
conventional accounts of the concerto genre. The opening movements of Moscheles’s
Concerto No. 7 in C minor (*Concerto pathétique*, 1835–6) and Bennett’s Concerto No. 4 in F
minor (1838) are cast in single-exposition concerto form, but have features that complicate
their relationship with the five types of Sonata Form set out in Hepokoski and Darcy’s
*Elements of Sonata Theory*. Both are similarly unusual in that they return to the tonic minor
for the second ritornello statement that by convention signifies the closure of the ‘larger
exposition’, and, probably as a consequence of the pre-emptory reappearance of Primary
material in the tonic here, their recapitulations are drastically abridged, complicating the post-
expositional stages of the form. These two works thus provide a rich theoretical and
hermeneutical challenge to the historiography of nineteenth-century instrumental form.
REFERENCES


Horton, Julian, 2007: ‘Sonata Form as Reception History: Schumann’s Piano Concerto and John Field’s Piano Concerto no. 7’, paper presented at the Sixth European Music Analysis Conference, Freiburg, 13 October.


The Musical World, 1/7 (29 April 1836).

The Musical World, 1/9 (13 May 1836).

The Musical World, 5/64 (2 June 1837).

The Musical World, 8/10 (8 March 1838).

The Musical World, 8/14 (5 April 1838).

The Musical World, 9/22 (31 May 1838).