Beyond the Creative Species

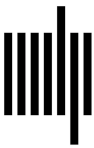
Making Machines That Make Art and Music

By: Oliver Bown

Citation:

Beyond the Creative Species: Making Machines That Make Art and Music
By: Oliver Bown
DOI: 10.7551/mitpress/10913.001.0001
ISBN (electronic): 9780262361750
Publisher: The MIT Press
Published: 2021

The open access edition of this book was made possible by generous funding from University of New South Wales.



The MIT Press

8 Speculative Futures

"Do you suppose there'll be a Third Industrial Revolution?"

Paul paused in his office doorway, "A third one? What would that be like?"

"I don't know exactly. The first and second ones must have been inconceivable at one time."

"To the people who were going to be replaced by machines, maybe. A third one, eh? In a way, I guess the third one's been going on for some time, if you mean thinking machines. That would be the third revolution, I guess—machines that devaluate human thinking. Some of the big computers like EPICAC do that all right, in specialized fields."

"Uh-huh," said Katharine thoughtfully. She rattled a pencil between her teeth. "First the muscle work, then the routine work, then, maybe, the real brainwork."

—Kurt Vonnegut[1]

So fast moving is this area, and so muddied by various forms of hype and emotional responses, that speculating about where we might end up even in a couple of years, let alone a decade or so, is a risky activity. But speculation is not prediction, and certainly not prophecy. It serves to point to and examine different possible futures, to allow discussion of where we want to go, what might need to be done to get there, and what we might want to look out for. I draw on the topics and examples discussed so far to analyze such possibilities and to frame the most prominent themes appearing in the emerging future of computational creativity. Most of the trends imagined below are already apparent and have already been identified in the preceding chapters, and are merely being extrapolated given what we can see the technology capable of, framed from the perspective of distributed, networked creativity. I divide this concluding chapter into three areas: industry, covering the nature of creative production and the use of commercially

developed computationally creative systems; society, covering the impact on cultural production and social behavior; and lastly, the most speculative and philosophical topic, the question of how machine creativity may cause us to reflect on our status as intelligent and creative beings.

In Industry

Speeding Up Creative Practice

AI is just one of a host of technological domains that has the potential to destroy jobs, threatening entire spheres of human work. Autonomous vehicles, if successful, have the potential to remove millions of paid driving tasks, not just professional drivers, but those for whom driving is a *part* of their work, and do so very quickly. Factory jobs are constantly being replaced. Now increasingly cognitive jobs are on the chopping board. A Bank of England report estimates that a full 50 percent of all current jobs could be at risk from automation. Some job losses are heavily resisted, as when the original Luddites went out and sabotaged machinery. But others slip away without much being said about them; photo development shops, once omnipresent, were obliterated by the advent of the digital camera, removing thousands of skilled and unskilled jobs.[2]

In a keynote on the impact of AI on jobs,[3] Toby Walsh argues that every business needs to have an AI plan, embracing AI technologies or mitigating the risks of AI innovation to their business. As a rule of thumb, he says, if you think of any human task that doesn't require much conscious thought, then it is likely that a machine can either do it or can just about do it now. For example, driving on freeways machines can do now, but driving in cities they can't. That might seem like a somewhat strange comment from the point of view of computational creativity; we effortlessly perform certain creative tasks which seem still far removed from machines' capabilities, but it nevertheless seems a reasonable rule of thumb in the majority of cases. Walsh points out that while Moore's law might have leveled off, algorithms continue to exponentially improve in performance, and funding for AI is doubling every two years. The momentum behind this transformation and its potential for exponential change is significant. But, he adds, while machine success in landmark competitive games, like Go and poker, came quicker than expected, the landscape of AI progress is bumpy and highly

unpredictable. Other predictions, such as those of Kurzweil[4] in computational creativity, have failed to materialize to plan.

Nevertheless, technology doesn't need to fully automate a task for it to erode jobs. Anything that helps someone do a task more efficiently or effectively—in other words, that increases productivity—will mean that fewer people are needed to get the same thing done. This has been happening in music, design, and media technology for some time through new powerful software tools.

At the same time, the overall social transformations caused by a technological development are harder to understand. Increased productivity doesn't always result in fewer people working in a sector. The reduced cost might mean that increasing demand takes up the slack or even drives up employment in a sector. This is, after all, what happens when a new technology first becomes viable. According to the Jevons paradox[5] in environmental economics, the increase in efficiency with which a resource can be used can lead to greater use of that resource. Consumption of coal increased after the invention of the Watt steam engine, which was a significantly more efficient engine than its predecessors. Wider cultural implications can also shift our approach to work. Judy Wajcman[6] argues that a current paradox of the acceleration of technology is that we have become more competitive and work focused, equating being busy with being successful.

The Stanford University report *Artificial Intelligence and Life in 2030*[7] summarizes the impact of AI on work as follows:

> AI will likely replace tasks rather than jobs in the near term, and will also create new kinds of jobs. But the new jobs that will emerge are harder to imagine in advance than the existing jobs that will likely be lost. Changes in employment usually happen gradually, often without a sharp transition, a trend likely to continue as AI slowly moves into the workplace. A spectrum of effects will emerge, ranging from small amounts of replacement or augmentation to complete replacement. For example, although most of a lawyer's job is not yet automated, AI applied to legal information extraction and topic modeling has automated parts of first-year lawyers' jobs. In the not too distant future, a diverse array of jobholders, from radiologists to truck drivers to gardeners, may be affected.
>
> AI may also influence the size and location of the workforce. Many organizations and institutions are large because they perform functions that can be scaled only by adding human labor, either "horizontally" across geographical areas or "vertically" in management hierarchies. As AI takes over many functions, scalability no longer implies large organizations.

Although the same report has a dedicated section on entertainment, it has little specific to say, except to list some of the applications of AI at present. So are creative jobs at threat from automation, and what changes are they likely to undergo? The above points, along with the themes I have built, frame some of the more convincing possibilities, which follow here. These following points, as I emphasize above, are speculations rather than predictions.

Certain domains of art practice will always place authorship and human connection front and center. In whatever ways these might be altered by automation, they will not take away the central role of the human as artist. A very large part of this domain of activity is amateur art production, where the individual achievement associated with the creation of the work is a critical factor, and the eradication of paid work is not a factor anyway (the economics of art is complicated by this vast reserve of potentially free human labor—people want to make art, with or without money involved). Although automation may transform specific activities it will not change the basic *use* of art as a human social pursuit based around individual authorship. In light of this, it seems unlikely that the machine-as-artist, an anthropomorphic agent with a personality and individual reputation, is likely to materialize outside of specific, highly conceptual instances.

Certain other domains of art practice are more faceless. Elevator music, advertising content, decorative art, building designs (outside of the realm of the superstar architect), and other similar areas all have the potential to be highly automated without disrupting artist-audience relations and other social patterns. What this means in practice, as outlined in the Stanford AI report, is that fewer people are needed to create the same work at the same rate, but there will still be individuals involved in the operation of such generative tools rather than full automation. This may be associated with a shift in skills. One possibility is that practitioners become spread over wider domains, a new generation of generalist creative producers: a single person produces an entire animated commercial including the storyboard, script, character animation, and music. Other more specific AI-related tasks might arise. Technicians specialized in the relevant technologies may emerge, as they have most prominently in the computer-focused creative industries such as animation, games, and areas of music. In architecture, we already have experts in building analytics who collaborate with experts in parametric design and multiobjective optimization. This could be a model for other areas.

In the extreme, major industries may rise up around specific areas of automated creativity. Examples include the automated generation of mood music by existing streaming services, content generation for interactive digital contexts such as games, and applications in targeted advertising. Generated or curated playlists on streaming services like Spotify are now a primary means by which people discover music. Some bring together highly author-associated works (like *Hits of the 70s*), whereas others are more faceless; they could already be generative without you knowing it. Context-based playlists on streaming services, such as workout or relaxation music, have become increasingly popular.[8] Tracks on such playlists are not strongly author-associated, yet generate serious income for their composers, and such companies have unprecedented power to control what music gets played on their platform via such suggestions. Elsewhere, large media organizations may create cultural products, such as AI pop stars, just as they already manufacture fictional characters and open-ended story worlds.

Certain production tasks such as audio mixing or mastering in the music domain are becoming automatable to the point that many types of content creators are already opting for automated solutions for cost reasons, even if the quality is diminished, just as we have seen in other areas (the canonical example being the replacement of live musicians with virtual instruments in orchestral screen composition). Other people may choose not to make this trade, and debates may remain unresolved indefinitely over whether human or machine outputs are better, just as questions of digital versus analogue audio production technology remain ongoing. Domains such as this might make interesting case studies for style arms races, where individual professionals attempt to differentiate their style from current machine capabilities. Nevertheless, it is plausible that the automation of tasks such as audio mixing and mastering has the potential to become acceptable and widely adopted very quickly, given that in machine learning terms the task is very clearly defined and exact input/output data is abundant.

In general, digital creative tasks may be sped up, but again due to any number of possible cultural factors this may not mean that creative production itself is sped up. It is possible that individual creators strive for greater quality or quantity given these powerful support tools. This may vary in detail from one domain to another. We may set the bar higher for quantity, or work might involve more background conceptual investment; the time it takes to make a conceptual artwork like Tracey Emin's *My Bed* may not be

reflected in the object itself. The conceptual development is invisible labor, often indistinguishable from other life activities. Alternatively, giving artists the potential for mass production through generative means might play out in different ways in different cultural scenarios: imagine an artist producing 7 billion artworks, one for each person on the planet, or composing 1,000 years of generative music (at least the latter, in a very simple sense,[9] and some variants of the former,[10] have already happened).

As well as speeding up creation, automation may shift the skill base. Being an expert in music theory or having the manual skill to draw real images may become increasingly less necessary in the capacity to create sophisticated works of music or art, as has already happened. Such automation may "democratize" creative tasks in a number of ways. As has happened in previous generations, newcomers with a new skill set and aesthetics may define new genres with their own cultural foundations and systems of value. The self-taught will be increasingly empowered with the help of technology, and yet they will seek new skill sets pertaining to the expert use of that technology. Content creation in general may become more of a participatory activity, at least in certain contexts, continuing the shift toward prosumer content creation associated with the widespread adoption of the Internet. In the spirit of Instagram filters, the idea that everyone can create *beautiful* things cheaply and easily can be taken forward in obvious ways by computational creativity technology. At the same time, establishment art forms may continue to uphold longer-standing skills, knowledge, and traditions.

Digital creativity is becoming increasingly networked and complex, associated with a number of transformations both technologically and culturally grounded. Creative practice in some domains may evolve toward richer and more complex ecosystems of practice. For example, forms of practice such as remix and mash-up, where digital artifacts are reconfigured, involve a greater exchange of content between individuals.[11] Creative coding tools such as MaxMSP and Processing emphasize reuse of code and libraries, meaning that creative practice becomes more cumulative, with the authorship of works involving more and more participants and complex chains of attribution. The rise of interactive media involves more numerous and more diverse creative collaborations between technologists and artists. With machine learning comes an additional pathway of attribution: when a neural network produces content we might trace attribution

to the programmer, the operator, the system itself, and in addition, the artists sampled in the training set. Here, software embodies aspects of cumulative culture, runaway niche construction[12] and the principle of increasing technological returns.[13] Layers are constantly being added to the ecosystem, and digital creativity can be said to have an incredibly diffuse authorship. As creative software capacities are added to this tapestry, they may perform subliminal roles but ones that only make sense from a distributed creativity perspective.

This final point is a restatement of the theme developed throughout this book that computationally creative technologies are already seeping gradually into creative practice in small niches within each creative domain, not in some explosive, disruptive transformation. The front line of commercialized computational creativity demonstrates this.

Starting in earnest around 2015, a number of startups have emerged in the UK, US, and Europe aiming to provide computer-generated music, and similar trends are seen in graphic design, advertising copy and concept development, design thinking, architecture, and video production. Some aim to sell music for high-budget synchronization to films, TV shows, and advertisements. Others aim to sell music in very low-cost scenarios such as the creation of amateur YouTube videos. Others still aim to sell generative co-creation tools to professionals.

The vast growth of amateur content creation means that millions of prosumers are in need of a jingle to place over their home video, and small companies are always looking for cheap music for product promos. My design students often produce concept videos for their design ideas, and as students, doing so with negligible production costs is critical to them. In all of these cases, the originality of the added music is not high on the list of criteria. Indeed, my students generally want to put their favorite music on their videos, copyright issues notwithstanding. The company JukeDeck is targeting this market, and aiming for a fully automated process. Another company, Landr, is providing automated mastering of tracks, with little in terms of options—simply submit your tracks and get them back mastered. Cultural studies researcher Jonathan Sterne's[14] research into Landr, whose business is now well established, shows that it is already transforming communities of practice. Mastering is a highly skilled process applied at the end of the music production pipeline but also a relatively expendable one. Amateur bands releasing music online might skip this stage altogether, and

many people would be hard pressed to tell the difference between a mastered and unmastered track, except that a mastered track will be more functional: it will fit in better on a playlist and sound better across different types of speakers. As Sterne notes, Landr's mastering algorithm may not stack up against a professional standard to the trained ear but is good enough for a large community of amateur producers to choose this much cheaper option, allowing them to rapidly produce and upload tracks that stand out more easily (in large part because the mastering makes the tracks sound louder). Interestingly, this includes musical fragments as well as finished tracks. Hip-hop beat-makers were seen to be using it on beats that they sold to rappers, a partly counterproductive approach, since traditionally the mastering would only apply after the vocals have been added (else there is no "room" in the mix for the vocals). Sterne also notes that a good mastering engineer may do very little to a track, but Landr always appears to be doing something, since doing nothing would appear a poor job (in the case of the real human, this may be a sign of great confidence instead).

Related to this, companies like Viddyoze aim to reduce the effort involved in making animated videos. They are doing this by templatizing animation creation, meaning that a user works from one of a range of templates, manipulating the properties that template provides. This could be seen as a form of casual creator system, packaging skilled creation tasks into a ready-made parametric system that can easily be personalized. The result is very limited in its versatility but nevertheless offers unskilled users the potential to create useful rich assets. Such templating methods are common in creative software, as in the selection of styles in Word or PowerPoint.

The startup Aiva uses a recurrent neural network trained on a classical music corpus to generate new musical scores. The music is generated in the form of a score for piano, which is then developed by the human composer into something for a larger ensemble. They target this use of neural networks at a specific stage in the workflow where they need to rapidly turn over initial proposals for a piece to show to a client, which the composer and client can agree on. Pierre Barreau, Aiva's CEO, says that this is the pain point he has identified in commercial music composition: there may be only a forty-eight-hour window in which to prepare something for a client. While it is possible to work at this speed, it doesn't guarantee the best results under time pressure. The goal of Aiva is not to replace the human

with a fully automated system but to support this process. The company is committed to the value of the human composer in the interaction with the client, counseling the client on the value of different approaches. But Barreau also points out the value of initially seeding the process with something automated, in that the composer can be more objective, supporting the client rather than being invested in one or other composition.

Aiva uses an out-of-copyright classical corpus to train its networks, meaning that there is no risk of a copyright issue arising. Nevertheless, they run a plagiarism checker on all output to ensure that no output ends up reproducing large chunks from the original corpus. But Barreau is confident that even such a simple learn-and-reproduce cycle can result in novel, creative outputs from the system. As discussed in chapter 5, he points to an example output from Aiva that sounds like traditional Irish music, despite being trained only on a classical corpus. This anecdotal example suggests that some properties of the music that transcend multiple styles may have been picked up, meaning that a system can generate output in a style despite not having been exposed to that style in the past. This is after all the principle of such a learning system—train it on ninety-nine works of a composer and it will in principle have the requisite musical structural knowledge to be able to produce an imaginary hundredth work.

The film and advertising sync world is an interesting target for an AI music business model; individual composers may be awarded highly lucrative royalty payments for very short pieces of content. It is a natural market to try to get into with computer-generated music. And yet, for the same reason it is already a market saturated by competition. Studios compete fiercely for those high-budget jobs, and every piece of music ever composed is a potential candidate for being picked up, with artists happy to get any such revenue for their music. Can there be a cost saving at the top end by automating this, as opposed to coming up with a simpler means by which existing music could be "rented" at low cost? This remains to be seen. But what is interesting in Aiva's case is that the focus is not merely on the cost of production but also on the specifics of the client experience. While we may not need more music, produced more quickly and at lower cost, we can still solve problems of how to help a client find the music they want more quickly. This is part of the business model for another company, Ecrett Music,[15] targeting the speed, accuracy, and satisfaction of discovery (not necessarily of production), *using* generative production.

As far as bringing automation to the world of commercial composition goes, then, a number of confounding factors distinguish this market from others. The most prominent is that, as already raised, creative music production is a desirable job. People have high intrinsic motivation to do it and a huge amount of amateur art-based work goes unpaid. It is hard even for a zero marginal-cost system to compete with that.[16] Also as already mentioned, while sync music is often faceless, there are times when the association with the artist matters; the musical meaning derives from the social connections that are involved in its production or dissemination.[17]

Copyright

Following such considerations, one pressing practical question surrounding the automation of creative production is the attribution of rights and the management of royalties covered by copyright law. The headline-grabbing question here is whether a machine might ever be granted rights that would otherwise be attributable to a human. If machines can be creative in the way that humans can, then do they deserve to be credited and rewarded? Hopefully the arguments in this book have already done enough to drive home the position that the capacity for the production of artifacts is one thing, but it is far from the total human experience, the holistic situatedness of people in the production of cultural artifacts. A system's ability to generate novel and valuable outputs is far from the only factor involved in considering rights attribution. Short of a *Blade Runner*–like future where robot replicants are agitating for recognition as people, validated as *total Turing test* winners (Hanard's T3, discussed in chapter 7), there is little sense in giving creative attribution to machines. Largely, existing copyright law already establishes this same position; in US copyright law, for example, a "human authorship requirement" is explicitly outlined, which rejects the possibility of attribution of copyright to machines and other nonhuman processes, including nature, animals, or plants.[18] This was recently upheld in the widely debated rejection of a copyright claim raised by activists on behalf of a monkey who took a selfie with a photographer's unattended camera.

The human authorship requirement still leaves open a number of possible outcomes for different human stakeholders. A content-generating AI system has associated with it various programmers and data operators, owners, and other individuals who contributed the content in the training data, all potentially laying claim to royalties and attributions from the systems

output. But this is not necessarily new; we already inhabit complex ecosystems of production. It is my view that introducing computationally creative systems, while certainly adding greater complexity and very likely introducing novel conventions of rights attribution, will not radically transform rights attribution as we known it and not diminish the logic of the human authorship requirement. Novel conventions of attribution are likely to arise as they do under any new mode of production. On a modern album, we may credit the mastering engineer for the album's sound but not the maker of the synthesizers or guitar amps, yet in some cases instrument-building might be more bespoke and warrant credit. Likewise, the programmer of a machine learning AI music system may be unlikely to warrant royalties for the generated music, but a creative user of the system is more likely to. Eric Drott[19] makes the case that copyright law is not well equipped to deal with the complexity (and ambiguity) of chains of attribution of machine-generated work, because it has so far been predicated on the idea that creation occurs, roughly, in individual brains, whereas AI systems are complex proxies for large groups of engineers and artists (through training data). McCormack, Gifford, and Hutchings[20] note that the disclosure of what data a system was trained on is a legal and moral gray area, but they contend that creators using machine learning–based systems have a responsibility to disclose what they trained those systems on.

A tantalizing argument is also emerging that machine-generated creative outputs should be placed in the public domain and not be copyrightable at all. If there is no immediate human author, only various stakeholders in the creation and operation of the algorithm, yet the algorithm is responsible for the creative innovation, then arguably its output would not (and should not) be covered by existing copyright law. In Japan this is being investigated with the question of whether such systems enable unfair competition.[21] If a machine is capable of bulk generation of cultural artifacts in a given style, then it could certainly have the potential to eclipse human creation in a disruptive and uncompetitive way. Does this matter? Do artists need protecting if their work in cultural production is being done for them? Such a debate will cut deep into the question of the social value of artistic creation and its cultural grounding. Bruce Gain draws attention to the way in which such machines are more like processors of existing cultural heritage than artists. "Copyright law is meant to protect people, not machines, and things created by machines should not be copyrightable by anybody.

They should be in the public domain," attorney Ray Beckerman says in an interview with Gain.[22] "There is already an excess of property right promotion in the US copyright world now, and not enough promotion of the public's rights in the public domain."

A counterargument is that AI generation is considered of positive cultural value and must be supported as an extension of human creative production. According to Andres Guadamuz, in an article for *The Conversation*:

> The most sensible move seems to follow those countries that grant copyright to the person who made the AI's operation possible, with the UK's model looking like the most efficient. This will ensure companies keep investing in the technology, safe in the knowledge they will reap the benefits. What happens when we start seriously debating whether computers should be given the status and rights of people is a whole other story.[23]

This is an interesting point because it highlights how at this intersection of art and technology we encounter an equivalent intersection between the respective logics of patent systems and creative copyright systems, which are quite different.

In another *Conversation* article, going further, Paresh Kathrani does actually make the case in favor of granting copyright to machines *themselves* on the basis that rights law should uphold intelligence in its own right, and not discriminate what the source of that intelligence is: "While machines may not be able to enforce intellectual property rights (yet), anything less might potentially amount to a violation of the value we place on intelligence in and of itself."[24]

This sits at odd with the primacy of copyright laws existing to protect people. If there is no apparent social need that is served by machines generating art and music, then why incentivize companies to create business models that do this? Whether artists need to be protected even when art gets made without their protection is a separate, long-standing argument that is itself politically charged, and is certainly beyond the scope of this book. The debates around computational creativity and copyright may only serve to reframe and perhaps even clarify such long-standing social issues.

However, note that removing protections for machine-generated creative works does not necessarily remove the power of commercial entities to disrupt and undermine existing forms of creative production. Drott reminds us that contemporary tech giants are successful because they are platforms—they create environments in which they can control and draw

rent from access or mine user data. Content and services may become free, but access to high-tech data services, such as of generation, becomes something that can be profited from. In other words, companies can make money even when they do not own the output of the work; indeed, making content free has been the path to great success.

As with all rules and norms, the shifting ground of copyright law is a product of struggles between different interest groups, which includes but is by no means exclusively limited to what objective observers would consider fair and socially beneficial. It is also enmeshed in society through feedback: if the rewards for creative production are greater, this may attract more people into creative areas, who then join the campaign to protect these rights.

Client-Producer Scenarios

Perhaps the most ready-to-hand metaphor of what interaction with future computational creativity systems might look like is a client-producer model, the seeds of which can be seen in the example of Aiva. According to this model, a creative practitioner requests a creative system to produce candidate outputs satisfying certain requirements. They may do this in any number of ways, using the paradigms discussed in chapter 6: direct manipulation of a neural network's training data and parameters; hand-coding a casual creator system; setting goals for a target-based evolutionary process; or having a system ambiently suggesting completions to what the user has started. The creative practitioner temporarily becomes client to the system-as-producer, expressing executive control, but not doing the work or necessarily exerting creative authority, since to do so means finer-grained control over the output—doing it oneself.

The client-producer model in its fullest manifestation involves a complete shift from users using software to build outputs to users requesting outputs from software (the request-based paradigm presented in chapter 6). Target-based and interactive evolutionary systems are a successful existing manifestation of this request-led approach, but their user scenarios still remain very limited for nonprogrammers and casual use.

One of the more tantalizing scenarios suggested in the previous chapter was an open-ended, service-based approach to computational creativity grounded in a natural language interface. It is easy to imagine a generative version of a search engine interface or language-based assistant. Just as you

type requests into the Google search bar or speak commands to an assistant such as Siri or Alexa, such a system would respond to natural language requests by parsing the request into a machine readable form and determining what actions to take to service the result. Rather than take the form of some general intelligence, it may be more plausible to envisage such a system resembling a generative web of subservices. Such a system might take a request and farm it out to a network of services: services for generating subcomponents such as melodies and rhythms; services for putting together higher-level structures such as verse-chorus arrangements; services for writing lyrics; style filter services; services for evaluating outputs to check for certain properties or to give a quality score; and services for setting up iterative processes bringing together generation and testing.

Furthermore, the possibility for such collaborations to be mutually adaptive exists through the potential to build systems that are tweaked and trained to work in the context of a specific user, like a personal assistant. The Stanford AI100 report says that "over the next fifteen years, the Study Panel expects an increasing focus on developing systems that are human-aware, meaning that they specifically model, and are specifically designed for, the characteristics of the people with whom they are meant to interact."

The suggestion of nonexperts becoming creatively empowered through computational tools is possibly a galling one for many creative professionals. As well as lacking the skills to create the content themselves, nonexpert clients may also lack an appreciation of the nuances, tropes, and stereotypes that distinguish the great from the feeble, the kitsch from cutting edge, the underlying mechanisms that make one font or color scheme more appropriate than another. Thus, automated processes may take hold for better or worse, eroding quality for convenience, just as virtual orchestral instruments have won over expensive orchestras in the production of most film music. But it is also possible that the affordances of such technologies instead improve and enhance creative output. Just as architects apply automated analytics to buildings, a client commissioning music could be given detailed analytics relating to their choices, just as music knowledge services have begun to provide tune analytics such as "danceability" and "hotness." Constraint solvers could ensure that generated outputs always achieve a set of requirements, such as legibility in print or audibility across a range of formats (sadly, in the world of music, this may include an even greater escalation of the *radio wars* battle to produce ever louder mixes of tracks).

Imagining a nonexpert user being both client and director, with a computational system supplying the domain-specific expertise, it is interesting to reflect on the agency relations here, the nature of the distributed agency, and the functional relationships between the various entities. Can such content-generation services act as instances of lofty computational creativity, and do they need to do so in order to fulfill their function? On the one hand, the director can assume the role of autonomous primary agent, controlling and directing the process, but as we have seen, a good service-provider system might be called upon to exhibit creative authority *over* the client. The musician Frank Zappa quipped that in the old days A&R people just wanted to know if the kids liked a band, whereas today they want to stamp their mark and make a legacy for themselves as curators. In this sense, we can imagine many different approaches taken by clients in relation to how much creative autonomy they allow computational systems, depending on the creative context.

New Domains of Interactive Creation

As well as ways of providing the automated generation of media, digital technology has created completely new forms of media experience that feed into a vision of how creative domains might evolve given computationally creative technologies. Digital environments are demanding increasing amounts of content that are generated on the fly, adaptive to context, or personalized. Computer games are leading this charge and, described as the "killer app" for computational creativity,[25] have drawn on all of the diverse generative creation scenarios we have considered in previous chapters. Yet generativity is emerging only slowly as a mainstream way to produce content. Good games are carefully designed to look and feel right. Generative algorithms are only trusted with certain tasks within narrow constraints, even if sometimes the lack of adaptive content creation can be jarring, as is the case with the predominant use of simple loops and switches in game music.

Meanwhile, games are really just the most obvious incarnation of a larger domain of interactive content, which is vast if you consider that it includes every piece of web design, the layout of every mobile app, every sound effect and transition animation. Any piece of media that is served digitally may not be interactive but has the potential to be, by virtue of the fact that it runs on a computational device in an interactive context.

Our daily experience is becoming increasingly augmented by digital media, with augmented reality at its frontier, which requires the production of media that is inherently adaptive, needing to adjust according to the context in which we are viewing it.

Another music-generation company, Melodrive, is interested in serving professional composers for games and other interactive experiences. Composing interactive music means creating systems rather than compositions. According to Anders-Petter Andersson, Birgitta Cappelen, and Fredrik Olofsson,[26] this can be thought of as designing strategies for decomposing and recomposing musical material: working out its composite elements and developing rules for how these elements can come together. As discussed in chapter 6, a major problem is that the various algorithms that have proven success in music generation in one area or another have not proven sufficiently generalizable in a use-case scenario. Currently the best ways to make generative music for interactive experiences in a way that enables the composer to have control over what they are doing are either very bland—the typical strategy is to cross-fade between different loops of music as required—or involve the composer to code their own systems. The latter is not unreasonable with the rise of the competent coder musician/artist, but even so, those practitioners benefit from any strategies and support tools that simplify or speed up the generative music creation process. Melodrive's approach is developer-focused, providing creative generative music algorithms accessible through programming libraries but also providing simpler workstation-based templates around specific themes. This assumes the rise of music developers to whom such services will be useful. These may be music-savvy developers or code-savvy musicians: new breeds of professional filling new technologically mediated niches.

In Society

Any impact on professional creative production will be paralleled by impacts on the experience of audiences (or "end-users," in the context of interaction design). What is computational creativity going to mean for consumers of art and music, as well as for those amateur producers who enjoy making it, outside of professional concerns?

The ability for people to generate astronomic amounts of creative content may seem an intimidating prospect, but really, we have lived for a long time in a world where no one person can experience even a small fraction

of all human-produced content. Even without automated means, merely through more routine developments that facilitate production (such as everyone having a video camera in their pocket), the quantity of creative content that exists is already incomprehensibly large.

Casual everyday creativity is as likely a target of generative technologies as any professional domain, given its immense scale.[27] The prosumer revolution has accentuated this great amassing of cultural products. It has arisen out of the basic conditions of the Internet and World Wide Web, that anything we create can be rapidly deployed to a public space where it can be consumed instantly by the rest of the world. We also interact digitally via private forums such as group and one-to-one chats, producing and consuming additional content in smaller communities. Most of this interaction is rapid and expressive, and the value of any technology that supports more expressive communication through these networked channels can be immense. One of music-culture theorist Jaques Attali's predictions[28] is that the world of music becomes increasingly many-to-many in nature, a more rapid-fire, fluid everyday process of creation. Although the prosumer revolution, where everyone with a smartphone becomes a public creator of literary and photographic aesthetic material, has partially driven such a scenario, we are still far from a situation in which creative art forms become casual media for everyday communication. Everyday language, photography, and video are probably the leading media for fluid communication but are generally distinguished from professionally produced movies, images, music, and text. Graphical art, music, animation, and more complex "professional" movie sequences are not commonly produced in daily interaction. Would generative creative tools, with their capacity to speed up creative production, change this? With the ability to employ machine tools to generate artworks and music, would these become increasingly used to play a role in everyday communication?

In one possible scenario, our daily, digitally mediated communication begins to take on richer and richer media forms, in much the same way that companies and entertainment producers use rich media to communicate to the public through video, audio, and increasingly interactive experiences. If it became trivially easy, and possibly also entertaining in its own right, to produce such content, then it is a relatively uncontroversial suggestion that we would do so daily for small audiences or even individuals. A party invite or a message to say "I love you" could take the form of

a unique, for-the-moment audiovisual spectacle. The use of looped GIFs, emojis, stickers, and memes in digital communication is the bedrock for this context. Messenger apps such as Snapchat and WhatsApp are increasingly facilitating the layering of media with such interactive and generative elements. Although these are relatively out-of-the-box functions, a shift toward increasing personalization would be a compelling development that could be served by generative tools. I think this domain of everyday communicative aesthetic interaction is where generative technologies will drive the most innovative developments of creative forms. Within this environment, artists and companies may find new participatory ways to interact with their audiences and customers, mediated by the support for creative expression enabled by generative technologies.

In this scenario, control-based, service-based, and ambient interaction paradigms are all plausible directions. For example, predictive text sets the scene for an ambient mode of generative interaction where suggestions for media are offered by the system according to what is being written. The user may be clearly active in the creation of the content (e.g., drawing shapes on a screen, or submitting source material to a style transfer algorithm) or may simply be selecting from suggestions proposed for the occasion (as we select animated GIFs, stickers, emojis, and meme imagery today). In turn, those suggestions may be generated by a system or may be more of a remixed or reappropriated nature, being drawn from a database of existing artifacts.

This obscures the creative autonomy not only of the system but of the user. The need for originality or virtuosity might be very low in such scenarios, and personal and group identity dynamics could dictate much of how creative roles are distributed. It is also plausible that we would produce such media for ourselves in an entertainment-on-demand scenario, playful interaction with a creative system as a means to an end, although it is hard to imagine people *not* sharing things—it is increasingly natural to casually drop one's creations on social media, even if they are produced in the course of casual entertainment.

Another dimension to this discussion is the rise of new aesthetic cultures that are, loosely put, characterized by "big data." Mash-ups and remixes are forms of cultural production that are premised on the iterative reworking of existing content. They are not necessarily original forms in this sense, and many art forms have rich histories of reuse, reference, and parody of existing content, but the evolution of mash-up and remix methods is underpinned

in part by new technological capabilities, beginning most notably with the audio sampler in music and Photoshop-like programs in visual domains. The sampler heralded the application of sampling techniques in hip-hop, other forms of urban music, and strands of art music in the late twentieth century. Now advanced information processing methods are continuing to drive new capabilities for mash-up and remix, as evidenced in the examples in chapters 5 and 6. Remix and mash-up practices are in turn related to the cultural shift from narrow cultural niches of taste toward a more "omnivore" approach to diverse aesthetic forms.[29]

Given that cultural production is laden with issues of situatedness, technologies like the sampler can be put to powerful use, as well as being tools of enhanced efficiency in their own right. Arguably, in addition, the digital audio sampler not only satisfied a need but also spurred on new creative forms through sheer experimentation and discovery that placed reference and reuse front and center. We could expect the same kinds of innovative use, adaptation, and application of emerging AI, evolutionary, and machine learning technologies, particularly in their application to big data.

Nefarious Uses

The election of the populist Donald Trump in the US was the culmination of a number of forces, one of which was the growth of social networks and their various risks and affordances. Analysis, including the claims of the campaign team technologists themselves, suggested that targeted memes of believable misinformation and emotionally salient materials were effective at mass manipulation, and that the dynamics of social networks provided new affordances by which social groups could form, consolidate, and potentially be controlled.

Social manipulation is, on this basis, a serious threat to society posed by AI. Cambridge Analytica, the company that performed data analytics via Facebook to support individually targeted campaigns for the Trump election and Brexit referendum, claimed to be able to produce detailed models of electorates.[30] Replacing statistical analysis based on focus groups and postcodes with the aggressive mining of online data using automated techniques leads to a completely different relationship between the public and those groups that wish to influence them. The entire population can be modeled and studied, and real-time trial and error of different strategies can be used to find election-winning formulae, resembling a highly effective

and creative evolutionary search process. The same techniques are available to companies, political parties, religious groups, and others whose intention is to influence and convert.

If the technologies of big data, social network analysis, psychological profiling and so on are proving effective at social manipulation, then the successful implementation of the technologies of computational creativity could add further capability to such efforts. Human art engages with such phenomena as the manipulation of emotion, the suggestion of meaning, the suspension of disbelief, the experience of the sublime, the attraction of attention, the creation of distraction, the creation of objects of desire, the stimulation of ecstatic and spiritual experiences, marking personal and group identity, and in the world of Internet epidemiology, the creation of highly shareable artifacts. Of particular relevance to this argument for the nefarious power of generative technologies is a body of research showing how the use of emotional, nonrational materials is at least as—if not significantly more— effective than rational argument, statistics, and appeals to expertise in effective social influence, in both product marketing and political persuasion.[31] But equally relevant are the themes developed in chapter 4 of art and music's potential function in social group formation and identity.

It is not inconceivable, then, that computational creativity researchers may find their research being used in scenarios removed from the playful creation of autonomous robot artists that help us reflect on human nature, and nearer to industrial-scale social manipulation.

One way that computational creativity could be used for nefarious use is in the creation of addictive activities like gambling that can be exploited for profit and can also be used to instill passivity in a population, limiting group coordination and political engagement. If cultural production supports culturally engaged individuals, then there is the potential that highly personalized automated cultural production creates feedback cycles that render individuals culturally detached. Another application could be in personalized emotional manipulation. This includes recruitment to a cause, inciting positive or negative emotions toward specific groups, manipulation of memory and associations, or tainting of experiences in real time. The creative generation of plausible misinformation, smears, and conspiracy theories is a well-established tool in political manipulation. We may imagine autonomous algorithms generating and testing possibilities directly on a population, without human evaluation or filtering. If one meme reaches

a million people, it hardly matters that 10,000 memes fail to go viral. Such techniques could be employed in the management of social groups. If we understand the dynamics of social groups and the role cultural production plays in their structure, stability, and evolution, then it may be possible to better force fission or fusion, enforce or erode cohesion and morale, stimulate in-fighting and so on.

Although some of these scenarios may sound farfetched, the concept that aesthetic objects influence people and may therefore steer them toward certain social affiliations is broadly supported by some of the work discussed in chapter 4. A danger lies in the potential of organizations to apply cultural influence on very large scales and over very long periods where tangible effects may be seen but are not obvious to those subject to them. Historically, attempts have been made to use music to motivate people in simpler ways. An example is the Muzak company, which experimented with the use of music in factories to stimulate greater work productivity. Their method was based on the concept of *stimulus progression*, where music is organized into blocks of time during which it gradually intensifies in tempo, loudness, and brightness. Although Muzak might have been claimed to have measurable effects, it was not a popular idea, and whatever benefits it might have offered were outweighed by competing factors. Musak lived on in the creation of "elevator music," with the company name becoming a generic term for bland mood music, which was considered functional in environments such as shopping malls where a mood was needed but an unobtrusive sonic profile was also considered desirable.

With linguistic creativity, the potential effects are more evident. Making up fake news stories that have meme-potential—plausibility and high salience—as well as being geared toward a certain political goal, can be put to powerful use. The extraction of highly salient details and biased, leading angles on otherwise factually accurate information, can also be applied to real news stories.

Even if not used in such purposefully nefarious ways, the rapid rise of industrial-scale AI applications has been highlighted as problematic in other ways, as in the analysis of AI and society researcher Kate Crawford.[32] Crawford notes how AI is often very dysfunctional in practice, even if it performs well in standardized metrics. This can be because the data used to train it is poor or simply highly ambiguous, especially when it lacks any depth of context. A significant problem arises when large complex

sociotechnical systems start to rely on such technology and to trust it when its trust hasn't been earned and without the necessary human support to fix resulting errors. In creative domains, poor training data—simple untagged corpuses of images and notes, for example—has the potential to misrepresent and dumb-down the complexity of cultural phenomena. Crawford also points to the power such industrial-scale AI systems concentrates in the hands of large corporations and their owners when they are provided as platform services.[33]

I have purposefully chosen to draw attention to some potentially very dark scenarios in the long-term application potential of creative technologies because, traditionally in computational creativity, the focus on social impact has largely been around how we would feel if machines created beautiful paintings and symphonies (see the following section), not how sociocultural dynamics will be impacted. This view stands in contrast to Smith and Leymarie's[34] optimistic vision of the impact of the machine as artist, and I feel it is an underrepresented position that needs to be considered more thoroughly. AI researchers and commentators have recently begun to mobilize around the dangers of AI, from the more obvious areas of automated weapons to nefarious uses in the financial industries. But given the great social importance of independent voices being politically and culturally autonomous, free of explicit or implicit influence, it is not too early to take a very serious look at the kinds of nefarious uses of creative technologies outlined in incredibly skeletal form here.

Beyond the Creative Species

Science writers have noted the historical "humiliations" that have come along with progress in the scientific understanding of the world. The earth is not at the center of the universe, or even of our own solar system, and humans are merely animals, evolved from a common ancestor with other life on this planet. We are neither central nor possessing of essential irreproducible qualities. This trend continues with the revelations of research in psychology, neuroscience, and evolutionary biology that reveals our biases, gene-driven motivations, failures of rationality and fast-brain (animal) responses. While psychology and neuroscience reduce the brain to something predictable, so AI threatens to demonstrate that machines can indeed do anything that we can, and in many cases already do so faster and better.

For as long as humans have existed we have been the most creative force we know of in the universe, but this seems soon to be placed into question.

Artistic creativity and artistic appreciation are, for some, central to this potential new phase of humility. David Cope, not confronted himself by such concerns, complains that for many people "it's all about machines versus humans, and 'aren't you taking away the last little thing we have left that we can call unique to human beings—creativity?'"—to which his response is "I just find this so laborious and uncreative." But how would people respond to evidence, not so much that *machines can be creative*—a question that, as I have discussed, invites enough ambiguity that we should just move past it—but that artistic behaviors can be modeled and predicted, that human aesthetic responses can be manipulated, and that creative production can be industrialized? What impact might we expect on the politics and social dynamics of identity, on cohesion and competition, on individual social relations and the socialization and enculturation of young people, and on the sense of satisfaction, enjoyment, and mutual appreciation that the creation of art and music provides for us, if we knew that all of these phenomena were modelable and predictable? Does computational creativity promise a new era of creative development and empowerment or will it herald, as the comment-stream contributor quoted at the very beginning of this book suggests, cultural redundancy and the loss of *everyone's souls*? This is a broad cluster of questions, and in this conclusion I will break the discussion down into two topics.

Machines, Souls, Authenticity, and Culture

A common complaint about machines making art is that they are merely *simulating* but are not *experiencing* as we do the aesthetic pleasure of art. They do not feel emotion and without this they cannot be legitimate creators of art. For some this failure may be manifest in poorer quality output; the simulation will always be incomplete and substandard. Machines will make refined rip-offs but will never produce sublime masterpieces or "truly original" work because to do so would require them to truly experience the emotion that we experience when presented with great art. For others, it may be that AI will be perfectly able to fake great creative work, but even if the work is in every way indistinguishable from human works of beauty, there would remain the conceptual issue that the machine is faking its experience.

With respect to the latter, the logic of the Turing test holds that we cannot pass judgment on what we cannot distinguish. Smith and Leymarie's optimistic appraisal is that we will, through developments in computational creativity, enter an era where machines understand the greatness of artworks as well as we do, an era of "humane machine intelligence," with "empathy and responsibility."[35] True to the spirit of Turing's test, if it *seems* empathetic and appreciative, then we must accept it as so.

The social perspective taken in this book skips this dilemma of true or simulated feeling, instead looking at what it means to be an art producer or consumer embedded in a cultural system, which I believe is what is really at issue when we consider what is the hard problem for efforts in lofty machine creativity. The algorithms used in computational creativity are largely derived from tools used to predict and to optimize, and insofar as there is data to predict or objective criteria to optimize, I believe that machines will be used to create amazing works of cultural production, following radically original paradigms. They will do so operating across the full spectrum from lowly generative machines, through complex collaborative co-creative interfaces, to lofty end-to-end creative agents that learn aesthetic principles, form their own conceptual spaces, perform search, and apply rules of novelty seeking and stylistic innovation. This is within the grasp of computational creativity, and systems that simulate or predict humans' emotional responses to aesthetic stimuli are also plausible.

Even if these systems perform beyond all expectations, the question of authenticity remains. From the social perspective this is not a matter of genuine feeling but of real situatedness. Just as with the philosophical question of authenticity of feeling, this is a question that lives outside of the technical, engineering domain, yet the social starting point is more tractable and empirically grounded than the philosophical one. As elaborated upon throughout chapters 3 and 4, humans apparently do not engage in art and music and other forms of cultural production merely as seekers of aesthetic enjoyment but as participants in processes of cohesion, competition, identity formation, and other social struggles. And there is evidence that the role of music and art in this effort is enmeshed not just in recent cultural history but in deep evolutionary history. For all their technical capacity, it will remain a social challenge to forcibly or otherwise embed an art- or music-generating algorithm into the role of either producer or consumer of art or music in ways that resemble our own cultural and biological situatedness.

This, for me, is a clear basis for rejecting the simple idea that through efforts in computational creativity we arrive at the creation of a "genuine machine artist," no matter what feats of creation that machine is capable of, not because it will never feel but because it will never sit in the situation of a person laden with a cultural history and life-determining choices about what creative actions to pursue. This is to take the view of "artist" as a social role. But the view of machine creativity presented in this book is one that is enacted through complex human machine networks. Within those networks, machines might fill many different roles and in their own way be culturally embedded systems. As their impact is felt and concepts of creative production are transformed, we may find new ways to think about what it means to be an artist, including extending the term to machines and transforming it as we do.

The social case against AI artists has been made by numerous commentators, but there is some difference in the detail. Sean Dorrance Kelly, a philosopher, argues that "human creative achievement, because of the way it is socially embedded, will not succumb to advances in artificial intelligence. To say otherwise is to misunderstand both what human beings are and what our creativity amounts to." While I agree broadly, I don't agree with the basis for this. He opines as I do that humans create artworks that reflect their situatedness in society, but for him it is insofar as this allows them to project a worldview:

> We count Schoenberg as a creative innovator not just because he managed to create a new way of composing music but because people could see in it a vision of what the world should be. Schoenberg's vision involved the spare, clean, efficient minimalism of modernity. His innovation was not just to find a new algorithm for composing music; it was to find a way of thinking about what music is that allows it to speak to what is needed now.
>
> … we must be able to interpret the work as responding that way. It would be a mistake to interpret a machine's composition as part of such a vision of the world.[36]

I see no reason why machines couldn't become adept independent social commentators, purveyors of visions and catchy or sociopolitically-relevant concepts, but the issue instead is where and why and with or for whom they perform such tasks; what is their grounding, and what skin do they have in the game of human social interaction?

New Creative Frontiers

Ahmed Elgammal, a key figure whose pioneering work on creative adversarial networks (CANs) we saw in chapter 5, navigates, as many commentators have done, between a vision of AI radically influencing the creation of art and of it having a comparable impact in many ways to previous transformations in the technology of art production. With Taishi Fukuyama he writes:

> Society is having an eerily similar debate to the one we had about photography and recorded or electronic sound back in the day. And the artists whose perspectives triumphed in those debates long ago because of the profound expressiveness of their art can speak to our concerns about AI. Machine learning systems and algorithms may prove not to disinherit us, but to become a new medium for human expression.[37]

But elsewhere Elgammal is keen to stress that AI systems are not simply tools, largely on the basis that they can be adaptive. Thus for him as with many creators of machine-generated works, it is important whether people accept the work "as art," asking if it is "able to qualify or count as art, and if it exhibits qualities that make it desirable or pleasurable to look at. In other words, could [CAN-generated] artifacts be recognized as quality aesthetic objects by human beings?"[38] This is a confounding issue: if the AI system is considered a mere tool, operated by an artist, then there should be little issue as to whether the output is art or not. So why should this be thrown into question? It arise as an issue only if the human operators were to attempt to render themselves invisible in the process. To recap, Gaut claims that art may be considered art for any of these reasons:

1. Possessing positive aesthetic properties
2. Being expressive of emotion
3. Being intellectually challenging
4. Being formally complex and coherent
5. Having the capacity to convey complex meanings
6. Exhibiting an individual point of view
7. Being original
8. Being an artifact or performance which is the product of a high degree of skill
9. Belonging to an established artistic form
10. Being the product of an intention to make a work of art[39]

The concern relates most clearly to item 10. In AI-generated artwork there is some intention to produce art, possessed not by the machine but by

the operator. For a machine intention to have any meaning, the operator (and their intention to make art) must be removed, *and* the machine must develop that intention independently; it can't be programmed to have the intention to make art! But since so many of Gaut's other criteria are trivially achieved, this is hardly a problem. The work can still be seen as art. The machine may be designed to predict human aesthetic responses, model cultural progressions, and optimize novelty within stylistic spaces, and perform very well at such tasks, producing legitimate art autonomously, as an author, but not an intentional one, not a being motivated as we are.

Thus while machine capabilities may cause us to take stock of our place in the universe, their application in the arts may do little to transform what role art plays in society, and will instead be adapted to those existing human motivations as technologies have already. AI may radically change how we *make* art, without necessarily replacing the primacy of the human artist. Whereas machines may drive cars and rockets and diagnose illnesses far beyond the limits of our intelligence and speed, the creation of art is measured as much by where humans are situated in social contexts as by the qualities of the content produced. Yet machines *may still* produce visual and sonic artifacts of great beauty and complexity, perhaps even finding stimuli that trigger emotional responses, awe, enchantment, and amazement beyond that previously known.

In conclusion, I return to the theme of function. In the view I have painted in this book, art and music are involved in the fluid social dynamics of cohesion and competition, with taste, style, and the desire for novelty acting as forces that shape this landscape. In some cultural contexts the production of art is widely dispersed and democratic, in others it is highly concentrated and stars are thrust to prominence through a winner-takes-all dynamic. In these various contexts the creation and exaltation of art can serve different individuals' goals. It may evolve in a way that is far beyond any one person's or group's conscious control, generative rather than adaptive in its nature, spawning new forms that do not exist for any particular purpose, which nobody could have predicted or even known they wanted. Increasingly, especially in music, this production has become more industrialized, serving corporate functions over social ones. Thus again I believe the greater risks associated with the automation of creative production are those associated with the way they might impact the overall shape and

dynamic of cultural production rather than the simple replacement of human activities. Such effects may take hold not only due to how good the algorithms are at generation but also how bad they are at fulfilling social roles in their entirety, a problem we see repeated elsewhere in the emergence of AI. We should enter this new world with enthusiasm for what may be possible but with eyes wide open to these broader potential impacts.

The open access edition of this book was made possible by generous funding from University of New South Wales.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Bown, Ollie, author.
Title: Beyond the creative species : making machines that make art and music / Oliver Bown.
Description: Cambridge, Massachusetts : The MIT Press, [2021] | Includes bibliographical references and index.
Identifiers: LCCN 2020017931 | ISBN 9780262045018 (hardcover)
Subjects: LCSH: Technology and the arts. | Arts—Technological innovations.
Classification: LCC NX180.T4 B68 2021 | DDC 700.1/05—dc23
LC record available at https://lccn.loc.gov/2020017931

10 9 8 7 6 5 4 3 2 1