Abstract—People with autism consistently demonstrate a lack of sensitivity to the full range of important aspects of everyday situations. Often, an overly restricted subset of the information available in a given situation gains control over their behavior. This can result in problems generalizing learned behaviors to novel situations. This phenomenon has been called overselectivity. Indeed, many behavioral intervention techniques seek to mitigate overselectivity effects in this population. In this paper, we offer an account of overselectivity as arising from an inability to flexibly adjust the attentional influences of the prefrontal cortex on behavior. We posit that dysfunctional dopamine interactions with the prefrontal cortex result in overly perseverative attention in people with autism. Limiting attention to only a few of the features of a situation hinders the learning of associations between the full range of relevant environmental properties and appropriate behavior. Thus, a restricted subset of features gain control over responding. A simple neurocomputational model of the attentional effects of prefrontal cortex on learning is presented, demonstrating how weak dopamine modulation of frontal areas can lead to overselectivity.

I. INTRODUCTION

Autism is a complex and varied developmental disorder affecting nearly 1 in every 150 people. The disorder is diagnosed by the presence of social impairments, communicative impairments, and repetitive/stereotypical behaviors and movements [1]. Numerous additional physiological and behavioral differences are commonly found in people with autism, making the task of identifying an underlying cause extremely difficult.

Some contemporary theories of autism have focused upon observed deficits in integrating contextual information in an appropriate manner. Problems integrating information, it is argued, result in a processing style which highlights the specific pieces of the environment at the cost of more general high-level information [2]. This “piecemeal” style of cognitive processing, described as Weak Central Coherence, is capable of explaining an impressive variety of both the advantageous and detrimental behavioral differences demonstrated by people with autism.

One possible mechanism that could give rise to information integration problems would be a difficulty in shifting attention from one perceived feature to another. If attention perseverates on an excessively small number of features, sufficient information to recognize more general properties of the situation remains unavailable. In this way, weak central coherence does not arise from an integration deficit, per se, but in an inability to fluently shift attention over all of the appropriate information to be integrated.

In addition to prompting a “piecemeal” processing style, attentional perseveration can also limit the degree to which relevant features of the environment are associated with particular behaviors. This failure to learn relevant associations has been called stimulus overselectivity, or simply overselectivity. Overselectivity can result in a reduced ability to generalize learned behaviors to novel settings. Failure to generalize is a major focus of many intervention techniques used in autism treatment [3], [4]. The role of overselectivity in limiting generalization is well illustrated by a study involving efforts to teach simple behaviors to children with autism [5]. Initially, each child was taught a new behavior (e.g., to raise their right arm when the phrase, “raise your right arm” was spoken) in one context. They were then moved to a new location, which included a new experimenter, and tested on their ability to generalize the recently learned behavior to the novel setting. Next, for those participants who failed to perform the trained behavior in the new context, items from the original setting were systematically introduced into the new situation. In many cases, the trained behavior could be elicited if some key feature of the original training context was reintroduced. Importantly, each of these participants appeared to be reliant on very specific, often idiosyncratic, pieces of the original training context. For instance, one individual required the exact same hand movements that were made by the original experimenter to be made by the new experimenter for any transfer to occur. Another child needed the table and chairs from the original room to be present before he would transfer the learned behavior to the novel environment. This experiment demonstrates how generalization can be problematic for people with autism, due to the learning of associations between the desired behavior and a restricted, possibly irrelevant, set of perceived contextual features.

Overselectivity in people with autism was first documented in the early 1970s by Lovaas and his colleagues. In this study, a compound stimulus including auditory, visual, and tactile components was presented to both low functioning children with autism and age matched control subjects [6]. (See Figure 1.) Initially, the subjects were trained to respond to...
the compound stimulus via an operant conditioning paradigm. Participants were rewarded when they made a specific action (e.g., a lever press) when the compound stimulus was presented. After the acquisition of this initial stimulus/response pairing, each individual component was presented separately to assess the degree to which the individual components, themselves, had acquired control of the behavior. In the normally developing control group, the participants responded equally to each of the individual components of the stimulus, demonstrating a lack of overselectivity. The group consisting of people with autism, however, responded to only one component of the three tested, thus demonstrating overselectivity. No systematic preference between the components was noted across subjects. This important result demonstrated how behavior in people with autism may be dominated by a small restricted feature set. This result has been replicated across various sensory modalities [7], [8] and using various numbers of features [9].

In this paper, we demonstrate that overselective behavior of the kind seen in people with autism can be explained in terms of dysfunctional interactions between the prefrontal cortex (PFC) and the mesolimbic dopamine (DA) system. Prefrontal cortex has been broadly associated with behavioral control, including the control of attention toward task-appropriate aspects of the current situation [10]. Computational models of PFC function have suggested that dopamine may play a central role in flexibly adjusting the PFC's attentional modulation of other brain areas as task contingencies change [11]. In these models, weakening the influence of DA on PFC can result in reduced attentional flexibility and a tendency to perseverate on only a few aspects of the current situation. Indeed, we have previously captured other behavioral patterns in autism, involving executive dysfunction, using a computational cognitive neuroscience model of perturbed DA/PFC interactions [12]. Using a similar computational model, we show in this report how the same kind of DA/PFC abnormalities can explain overselectivity in autism.

II. PFC/DA Interactions

Our previous modeling efforts have demonstrated how impaired interactions between the DA system and the PFC can result in perseverative attention to a restricted set of stimulus dimensions and features. In particular, we showed how modifying the effect of DA on the PFC can produce deficits on the Wisconsin Card Sort Task (WCST) that match those seen in autism [12]. These effects arise naturally in our computational model of frontal function, which has been shown to capture human performance on a variety of executive tasks as conducted by both patients with frontal brain damage and healthy controls [13].

Our account of PFC function was developed primarily to explain the role of PFC in cognitive control and cognitive flexibility. Cognitive control is the ability to guide behavior according to explicit goals or rules, especially when doing so is in conflict with more automatic or prepotent tendencies. Cognitive flexibility describes the ability to appropriately adapt cognitive control in response to shifting task contingencies. The PFC has been broadly implicated in cognitive control and cognitive flexibility [14], [15]. In our models, the PFC supports cognitive control by actively maintaining abstract rule-like representations that provide top-down modulation of more posterior brain areas, modifying the regular behavior of these posterior pathways so as to overcome their usual automatic patterns of responding [10]. Biologically, the active maintenance of frontal control representations is supported by dense patterns of recurrent excitation in the PFC, as well as intrinsic maintenance currents [16]. Computational models of these neural circuits have shown that the active maintenance of control representations and the flexible adaptation of control are at odds, with the mechanisms that maintain PFC representations, and protect them from distracting inputs, acting as an obstacle to the rapid updating of PFC contents in response to shifting contingencies. Thus, in order to achieve cognitive flexibility, a separate mechanism is needed to intelligently and rapidly update the actively maintained PFC control representations in a task-appropriate manner.

A useful analogy for this flexible updating mechanism is that of a “gate” in a fenced enclosure. When cognitive
control must be strong, the gate is closed, keeping out distracting inputs that might compromise the needed PFC control signals. When the current control state is no longer appropriate, the gate opens, allowing the old control state to escape and allowing a new control representation to enter the PFC via its inputs. What is needed is a neural mechanism that can adaptively open and close this gate on PFC in a task-appropriate manner. Some researchers have suggested that the mesolimbic dopamine system may play a central role in controlling this gate [11]. These cells have been found to carry reward prediction information critical for learning associations between behaviors and reward [17], and the DA projections to PFC have been viewed as a likely neural implementation of the gating signal needed to flexibly adjust the control state of PFC [11].

Under this account, DA interactions with PFC drive the flexible updating of control. Inflexibility arises when these interactions are disturbed, frequently resulting in PFC perseverating on control representations that are no longer appropriate. This insight, along with evidence of DA abnormalities in autism, has led us to investigate the degree to which the perturbation of DA/PFC interactions naturally leads to patterns of behavior observed in people with autism. Our previous computational modeling work has shown that this mechanism is sufficient to explain various aspects of executive dysfunction, including Stroop and WCST data, in autism [12]. In this paper, a slightly more abstract computational model of this mechanism is used to show that perseveration of PFC control representations, in this case controlling attention to stimulus features, brought about by disturbed DA/PFC interactions, results in overselectivity. Thus, in addition to accounting for executive dysfunction, our account is shown to also capture the pattern of stimulus overselectivity seen in autism.

III. Modeling Overselectivity

In the conditioning paradigm used by Lovaas and his colleagues to assess stimulus overselectivity, the severity of overselective behavior is measured by noting the number of the compound components capable of eliciting a response in isolation from the others. Overselectivity occurs when the number of components capable of driving a response in isolation is lower than the total number which comprise the compound stimuli. If an individual responds to all components equally, this indicates that attention has been distributed across all components during learning and no overselectivity is demonstrated.

A simple artificial neural network model, constructed using the biologically grounded Leabra framework [18], is utilized here to model overselectivity phenomena. In this model (see Figure 2), an input layer represents the presented compound stimulus. Each unit of this layer reflects an individual component of the stimulus. For example, the first three units can be thought of as representing an auditory, visual, and tactile component, respectively. To represent the compound, all three individual units are clamped to a high activation value, simultaneously. The Hidden Layer learns stimulus to response mappings, and provides a modeled abstraction of posterior brain systems. A Response Layer encodes the output of the network. The two possible outputs represented within the Response Layer of the network are “Respond” and “Do-Not-Respond”, with strong lateral inhibition between these units encouraging unambiguous decisions. Additionally, a PFC layer provides a top-down influence on processing within the Hidden Layer (posterior cortex). The PFC layer has one extra unit which is utilized in the working memory load condition described below. (Note, however, that this working memory load unit is not shown in Figure 2.). The input layer also contains one extra unit, and this unit represents a “No Stimulus” condition, analogous to when the compound stimulus is not being presented. Each unit in the PFC layer is associated with exactly one unit in the input layer. Each of these input/PFC pairs project to a unique pool of Hidden Layer units, producing an isolated processing pathway for each stimulus component, modulated by its corresponding PFC unit. This enables each PFC unit to have selective influence upon a unique component of the compound stimulus. In other words, each hidden unit directly participates in only one of the three possible pathways. Note, however, that there is lateral excitatory connectivity between all of the units in the Hidden Layer, with the strength of these connections determined during conditioning by Leabra’s learning mechanism. Also, as is standard in Leabra, the Hidden Layer incorporates fast pooled inhibition, limiting the amount of activity across the processing units in the layer. Thus, one processing pathway can possibly influence the computations performed by another pathway through these lateral interactions.

In order to model further demands placed on PFC in the form of a working memory load, a fourth unit is included in the PFC layer. This unit selectively projects to an additional isolated pool of hidden units. These hidden units are also fully recurrent, and thus connected to all of the other Hidden Layer units, as described previously. Only one difference exists between the working memory load units and the units associated with the stimulus components. Namely, the hidden units receiving input from the working memory load PFC unit do not receive any projections from the input layer, as the working memory load is seen as supplemental to the processing of the compound stimulus.

When modeling both normal and autistic performance, the network is conditioned to “Respond” to the compound stimulus by repeatedly activating the three stimulus input units simultaneously, representing an auditory, tactile, and visual component, respectively, and providing error feedback on the produced output. With each stimulus presentation, connection
weights are modified using the standard Leabra learning algorithm, incorporating both error-driven and correlational mechanisms [18]. This is an extremely easy task for the network to learn, as it only needs to associate a response with the stimulus that is presented. This duplicates the simplicity of the original behavioral study by Lovaas and his colleagues. In order to model healthy PFC function, in which attentional control is flexibly adjusted during learning, activity in the PFC layer is allowed to switch between all three possible states, attending sequentially to each of the three stimulus components. When modeling autism, however, only a single arbitrary unit of the PFC layer is activated, and this unit remains active throughout the entirety of training. Perseveration of the PFC control representation on a single stimulus feature simulates a deficit in flexibly updating PFC contents. This difference in the ability to flexibly adapt representations actively maintained by the PFC is the only difference between the model of normal performance and the model of autistic performance. All other parameters are identical between the two models. After the initial stimulus/response training with the compound stimulus, each component is presented individually to the network and the network’s output is recorded. During these single-component testing trials, PFC activity is set to attend to the single presented stimulus feature. This PFC state is used in both the healthy and autistic networks, avoiding situations in which attention is directed to a feature that is not actually present in the stimulus. The measure of interest is the number of individual components capable of correctly producing a “Respond” output from the network.

Additional support for this PFC-based model of overselectivity can be found in an intriguing recent study which suggests that stimulus overselectivity can be induced in healthy individuals by requiring the concurrent performance of a working memory task [19]. Working memory tasks are widely believed to enlist the resources of PFC, providing support for the conjecture that healthy individuals utilize this area when learning to associate a compound stimulus with a response. In order to investigate whether our model can capture these results, an irrelevant additional item can be maintained in the PFC layer during the learning phase, simulating a working memory load. This is achieved by keeping an extra PFC unit constantly active throughout the entirety of training, simulating maintenance of extra information within the PFC. All other parameters in this working memory load condition are identical to the control condition. This includes the flexible updating of the PFC, which is allowed to flexibly switch attention over the three stimulus features during training.

IV. RESULTS

Network simulations were repeated 100 times in each of the experimental conditions, with initial synaptic weights randomized for each repetition. Average performance results for each condition were compared. These simulation results qualitatively match human performance, providing evidence that a lack of flexibility in PFC updating can result in a restricted cue set gaining control over behavior. (See Figure 3.)

The model of autistic performance responded to significantly fewer components (\( p < 0.05 \)) compared to the model allowed to flexibly update its PFC representations, demonstrating overselective behavior in the autism model. Providing additional support for the hypothesis that the PFC influences learning in other cortical areas in interesting and important ways, a modeled working memory load during training of a healthy network resulted in significantly more overselective responding (\( p < 0.05 \)), mirroring recent behavioral results.

Overselectivity arises in this model due an effect of PFC-directed attention on the learning of associations between stimulus features and the response output. When PFC activity is directed to a particular stimulus pathway, the activation lev-
els of the hidden units of that pathway are increased. Lateral inhibition within the Hidden Layer, driven by this increased activity, reduces activity in the pathways corresponding to the other stimulus components. Thus, learning primarily takes place within the selected pathway, as synaptic plasticity is strongest in Leabra in the presence of presynaptic activity. In the autism network, which remains focused on a single pathway throughout training, the synaptic weights grow strong only within the selected pathway, with the connections in the other pathways remaining relatively weak. Thus, the autism network fails to learn to generate responses to the unattended stimulus features, even if attention is later directed to those features. In contrast, the healthy model flexibly adjusts PFC activity during training, allowing different pathways to be strengthened on different trials, eventually producing strong associations between all of the stimulus features and the need to respond. Leabra’s Hebbian learning mechanism further ensures that these associations are robust.

V. DISCUSSION

The breadth of neurological abnormalities discovered in people with autism is almost staggering. The neuropathology of autism includes, but is not necessarily limited to, the cerebellum, PFC, anterior cingulate, hippocampus, amygdala, temporal lobes, parietal lobes, and various neurotransmitter systems [20]–[24]. Given the ample choices of brain areas to investigate, many with ties to behavior in autism, why champion a closer look at the role of dopamine and its interactions with PFC?

Neurally, dopamine affects nearly all of the brain areas associated with autistic behavior [25]–[30]. The evidence for DA abnormalities in autism is strong. Both PET imaging studies and urinalysis studies reveal differences in levels of DA in people with autism [31], [32].

Dopamine plays a major role in many of the problematic behaviors demonstrated by people with autism. These behaviors range from non-intentional behaviors, such as seizures, to those at a much more cognitive level, such as learning to follow eye gaze and the ability to control and flexibly adapt our behavior. The breadth of these links is, perhaps, the strongest argument for a closer examination of the causal role of DA in autism. Approximately 1 in 4 people with a diagnosis of autism will develop seizures during adolescence, significantly higher than the 1 in 200 prevalence observed in the general population [33]. Researchers believe that DA plays a major role in epileptic seizures [34], and anti-convulsant medications are known to have direct affects on the DA system.

People with autism also demonstrate a variety of motor difficulties, including problems initiating behaviors, repetitive movements, and abnormal gaits [1], [35], [36]. The basal ganglia and the mesolimbic DA system are widely accepted as a vital component in learning and initiating motor movements. Problems within these brain areas are believed to be at the root of disorders with substantial motor control symptoms, such as Parkinson’s and Huntington’s disease. Also, stimulation of the DA receptors located within the striatum has been shown to induce motor stereotypies, and these are ameliorated or abolished by blocking DA transmission within the striatum [37]–[39]. Thus, even stereotypic behavior, one of the triad of impairments currently needed for an autism diagnosis, has direct ties to DA.

The ties of DA to autism are both numerous and compelling. Further investigations into precisely how these differences affect behavior has great potential for providing a common language, that of neurobiology, for linking many disparate behaviors in people with autism.

VI. CONCLUSION

Since Kanner’s original description of “early infantile autism” in 1943, it has been noted that people with autism seem to be preoccupied with specific and sometimes peculiar parts of objects and situations [40]. This frequently leads to stimulus overselectivity and the concomitant failure to generalize learned behaviors to novel contexts. Generalization of this kind is central to daily functioning, so serious difficulties in this area can be crippling.

The modeling results presented here suggest that, in people with autism, overselectivity may be driven largely by abnormalities in DA/PFC interactions, causing inflexibility in the shifting of top-down attention. When the PFC is unable to flexibly and appropriately update its contents, representations in cortical areas downstream from the PFC develop which are dominated by an overly restricted, or possibly even irrelevant, subset of features from the environment. Poor generalization occurs, under this account, due to the same abnormal cortical representations. The inability to flexibly update the PFC increases the likelihood that the only environmental associations that will be learned in a given situation will involve spurious correlations (e.g., idiosyncratic features of the training process), with other, more relevant, factors escaping attention. Subsequent dependence on such spurious correlations can cripple generalization performance, as described in the introduction. The results presented here, coupled with past research on how DA/PFC impairments can explain executive dysfunction in autism, provide support for a common neurological cause underlying a variety of behaviors observed in autism. Indeed, ongoing work is investigating the role of DA/PFC interactions in information integration problems in people with autism. It is possible that the presence of a DA/PFC impairment over extended developmental timescales may lead to behavior which looks like an integration problem on the surface, but is actually just integrating an overselected range of information.

VII. ACKNOWLEDGMENTS

The authors extend their gratitude to the members of the Computational Cognitive Neuroscience Laboratory at the
University of California, Merced.

REFERENCES


