

## O PAPEL DO CONTEXTO NA PERCEPÇÃO DE EMOÇÕES<sup>1</sup>

---

Felipe Nogueira de Carvalho<sup>2</sup>

### RESUMO

De todos os aspectos do comportamento não-verbal, a face é sem dúvida uma das mais ricas e importantes fontes de informação sobre o estado interno do outro. Mas expressões faciais são raramente percebidas de forma isolada. Ao contrário, são tipicamente inseridas em contextos sociais ricos e dinâmicos, que incluem gestos e posturas corporais, conhecimento situacional, etc. Com base nessas observações, podemos nos perguntar se o contexto no qual uma expressão é percebida pode influenciar a percepção de emoções nesta expressão. No caso de uma resposta afirmativa, de que modo se daria essa influência contextual, e quais seriam os seus limites? O propósito desse artigo é explorar algumas possibilidades sobre o papel do contexto na percepção de emoções, desde a teoria das emoções básicas, que defende que categorias discretas de emoções podem ser lidas diretamente da face de forma invariável, a abordagens mais contemporâneas, que atribuem um papel constitutivo para o contexto na percepção de emoções. Embora o debate esteja longe de ser resolvido, as conclusões deste artigo apontam para um novo modo de se pensar sobre fenômenos emocionais, onde a díade de interação torna-se a unidade básica de análise, e onde emoções são concebidas como propriedades emergentes de relações em contextos particulares de interação social.

**Palavras-chave:** Emoções. Percepção de Emoções. Expressões Faciais. Cognição Social.

### ABSTRACT

From all aspects of non-verbal behavior, the face is undoubtedly one of the richest and most important sources of information about the internal states of others. But facial expressions are rarely perceived in isolation. On the contrary, they are embedded in rich, dynamic social contexts that include body gestures and postures, situational knowledge, and so on. On the basis of these observations, we can naturally wonder if the overall context in which the face is embedded can change how emotions are perceived in facial expressions. If so, in what ways, and what are the limits of these contextual

---

<sup>1</sup> O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. Agradeço a Ernesto Perini e Carlos Barth por comentários e sugestões feitos a uma versão anterior desse artigo.

<sup>2</sup> Pesquisador de pós-doutorado no departamento de Filosofia da UFMG pelo programa CAPES-Print. E-mail: [felipencarvalho@gmail.com](mailto:felipencarvalho@gmail.com).

effects? The purpose of this paper is to explore some possibilities concerning the role of context in emotion perception, from basic emotion theory, who holds that discrete emotion categories can be read out directly from the face in an invariant manner, to more contemporary approaches that assign a constitutive role for context in emotion perception. Although the debate is far from settled, the conclusions of the paper will point to a new way of looking at emotional phenomena, where the dyad of interaction becomes the basic unit of analysis, and where emotions are conceived as emergent properties of a relationship in particular contexts of interaction.

**Keywords:** Emotions. Emotion Perception. Facial Expressions. Social Cognition.

## I. Introdução

O comportamento não-verbal é muitas vezes tão essencial à nossa navegação diária do mundo social quanto a comunicação linguística. O sorriso radiante de um ente querido após a recepção de um esperado presente de aniversário, ou o olhar penetrante de um chefe quando chegamos atrasados no trabalho, podem muitas vezes transmitir mais do que palavras, nos ajudando a compreender o estado emocional do outro e a ajustar nosso comportamento verbal e não-verbal de acordo com a informação emocional obtida.

De todos os aspectos do comportamento não-verbal, a face é sem dúvida uma das mais ricas e importantes fontes de informação sobre o estado emocional do outro. Bebês, por exemplo, começam a perceber e a demonstrar preferências por rostos desde os primeiros dias de vida, e há evidências de que já conseguem discriminar expressões positivas e negativas desde essa idade (FARRONI et al. 2007). À medida que nosso sistema cognitivo amadurece, continuamos a nos apoiar em configurações significativas de musculatura facial como guias confiáveis para o estado emocional do outro. Não é, portanto, de admirar que o estudo das emoções ao longo do século XX tenha se pautado em grande parte por expressões faciais.<sup>3</sup>

Mas expressões faciais são raramente percebidas de forma isolada. Ao contrário, são tipicamente inseridas em contextos sociais ricos e dinâmicos, que incluem posturas e gestos corporais, tom de voz, informação situacional, e assim por diante. Se possuímos a informação colateral de que

---

<sup>3</sup> Para um apanhado histórico ver Gendron & Barrett 2009.

Lúcia está com dor de estômago, podemos perceber seus movimentos faciais em termos de dor e desconforto; mas se sabemos que ela está se lembrando de uma experiência particularmente embaraçosa, os mesmos movimentos poderão ser percebidos em termos de vergonha e embaraço.

Com base nestas observações, podemos nos perguntar se o contexto em que uma expressão é percebida pode modular a percepção de emoção nesta expressão. No caso de uma resposta afirmativa, a que tipo de informação contextual a percepção de emoções seria sensível, e quais seriam os limites de tais efeitos contextuais? No caso de uma resposta negativa, como explicar a variabilidade na expressão e reconhecimento de emoções em diferentes contextos? Qual é, afinal de contas, o papel do contexto na percepção de emoções?

Estas são questões altamente controversas no campo de estudo das emoções. Alguns teóricos defendem que categorias emocionais podem ser lidas diretamente da face de forma automática e invariável, enquanto outros admitem uma ampla gama de efeitos contextuais sobre a percepção de emoções. Outros foram ainda mais longe e negaram a própria ideia de “percepção de emoções”, visto que expressões emocionais não seriam objetos pré-definidos, mas propriedades emergentes de relações sociais.

O propósito desse artigo é explorar algumas possibilidades sobre o papel do contexto na percepção de emoções. Começarei analisando a teoria das emoções básicas, que defende que expressões faciais seriam automaticamente produzidas e reconhecidas por mecanismos biológicos dedicados de forma invariável (EKMAN 1992). Variações contextuais entrariam apenas após a conclusão desse processo, através de convenções, normas e hábitos de uma dada cultura que dizem respeito ao gerenciamento de respostas emocionais. Na seção III examinarei em seguida a teoria dimensional de Russell (1997), que confere um papel importante ao contexto sob a forma de conhecimento situacional. Nesta teoria, expressões faciais transmitem informações sobre valência e excitação de forma invariável, mas a classificação de expressões em categorias particulares de emoções é um processo inferencial sensível ao contexto.

Este modelo foi questionado a partir de evidências de que mesmo a percepção de valência e excitação parece estar sujeita a fatores contextuais.

Com base nessas evidências, irei discutir na seção IV o que chamarei de “modelo da percepção de objetos”, que lida com a percepção de expressões faciais de forma similar à percepção de objetos materiais comuns. Esse modelo, no entanto, será criticado por trabalhar com estímulos estáticos sem validade ecológica, já que o comportamento facial e corporal é um processo dinâmico em constante mudança à medida que a interação se desdobra no tempo. Para acomodar essas observações irei apresentar em seguida o modelo da sincronia conceitual (seção V), onde a percepção de emoções será reformulada como a co-construção de emoções por dois parceiros afetivamente engajados em uma interação social.

Essa discussão abrirá caminho para um novo modo de se pensar em fenômenos emocionais (seção VI), onde a díade de interação torna-se a unidade básica de análise, e onde emoções são concebidas como propriedades emergentes de uma relação em contextos particulares de interação social. Nesta visão, o contexto desempenha um papel constitutivo na percepção de emoções, visto que emoções não podem ser compreendidas fora das dinâmicas particulares de interações sociais. Finalmente, a seção VII esboçará algumas conclusões.

## II. A Teoria das Emoções Básicas

Em sua obra “*The expression of the emotions in man and animals*” (2005/1872), Charles Darwin propôs que expressões faciais teriam evoluído para ajudar a preparar o organismo para responder adaptativamente a certas situações ambientais recorrentes. Quando o organismo está diante de uma situação ameaçadora, por exemplo, os olhos se arregalam para aumentar o escopo do campo visual, resultando na expressão característica de medo que conhecemos. Os movimentos faciais que, de forma conjunta, constituem a resposta de medo, tornam-se aqui uma “macro” comportamental adaptativa, juntamente com outras respostas corporais como batimentos cardíacos acelerados para redistribuir o sangue em preparação para movimentos rápidos, ereção capilar para o organismo parecer maior para um predador, e assim por diante.<sup>4</sup>

---

<sup>4</sup> Para uma discussão crítica das ideias de Darwin sobre expressões faciais, ver Barrett 2011.

Inspirada pelas observações de Darwin, a teoria das emoções básicas propôs uma lista de emoções primárias como respostas automáticas a situações ambientais recorrentes, envolvendo temas universais como ameaça, perda, etc. Seriam estas: medo, raiva, tristeza, repulsa, surpresa e contentamento (EKMAN 1992).<sup>5</sup> Supõe-se que essas respostas sejam instanciadas por mecanismos biológicos dedicados, geneticamente programados para ativar sempre que o organismo se deparar com certas situações recorrentes que ajudaram a selecionar essas respostas no passado.<sup>6</sup> Quando interações sociais tornaram-se mais proeminentes na história evolutiva do organismo, estas expressões passaram a servir uma função importante de comunicação de informação social, permitindo que indivíduos co-específicos fossem alertados sobre a presença de certos estímulos apenas pela percepção de certas configurações faciais em outros.

Com o tempo estas expressões tornaram-se ritualizadas, isto é, mais exageradas e prototípicas, e foram selecionadas neste formato por permitir um reconhecimento mais rápido e menos ambíguo da parte de co-específicos (SHARIFF & TRACY, 2011). Desta forma, mesmo que um organismo não precise aumentar o escopo de seu campo visual em uma dada situação de ameaça, seus olhos ainda assim se arregalariam como parte de uma expressão ritualizada de medo, selecionada por sua função de comunicar efetivamente o perigo iminente. Nesta teoria, a produção automática de certas expressões faciais na presença de situações ambientais recorrentes, e o reconhecimento automático destas expressões como transmitindo certos tipos de informação social, são ambos produtos da evolução (*ibid.*, p. 397).

A evidência empírica para a teoria das emoções básicas veio através de estudos interculturais independentemente conduzidos por Paul Ekman e Carrol Izard nos anos 60 e 70.<sup>7</sup> Pois se a teoria das emoções básicas está correta, e expressões faciais são respostas automáticas selecionadas pela evolu-

---

<sup>5</sup> Outros teóricos de emoções básicas operam com categorias ligeiramente diferentes, mas a extensão exata da lista não é importante para os propósitos deste artigo. Para uma discussão informativa ver Crivelli & Fridlund, 2019.

<sup>6</sup> Qual estímulo particular irá ativar uma resposta emocional depende do aprendizado e, portanto exibe um alto grau de diferença cultural e individual (embora respostas de medo a certos estímulos, como serpentes, possam ser inatos; ver Ledoux 1998). Mas uma vez que um estímulo seja avaliado como ameaçador, ele irá automaticamente ativar respostas de medo a cada encontro subsequente.

<sup>7</sup> Essa evidência é descrita e revisada em Ekman, 1989.

ção a situações ambientais recorrentes, então estas expressões devem ser universais, e automaticamente reconhecidas por indivíduos de diferentes culturas como expressões pertencentes a uma mesma categoria básica. Em um experimento típico, apresenta-se a um participante descrições curtas de situações (“um parente próximo faleceu, e ela está triste”), e o participante deve escolher, a partir de uma série de fotografias representando expressões prototípicas de emoções básicas, a que melhor captura o que a personagem da situação deve estar sentindo.<sup>8</sup> Esses estudos foram amplamente replicados em mais de uma dúzia de culturas ocidentais e não-ocidentais, incluindo uma cultura pré-alfabetizada e visualmente isolada da Nova Guiné, e os resultados apontaram para uma evidência esmagadora de universalidade, sob a forma de um alto nível de acordo inter-cultural sobre a expressão que melhor combina com cada história apresentada (EKMAN, 1989).

Baseados nesta evidência, teóricos das emoções básicas propuseram a “hipótese da leitura” (*readout hypothesis*), segundo a qual categorias básicas podem ser lidas diretamente da face de modo puramente *bottom-up* (Buck 1994). Mesmo que outras informações corporais, como os punhos cerrados de raiva, sejam igualmente parte da expressão emocional, categorias básicas podem ser automaticamente reconhecidas até mesmo em uma imagem instantânea da face apresentada de forma isolada, como afirmado por Ekman:

As informações que [expressões faciais] transmitem sobre emoções podem ser capturadas em um único instante. Tipicamente, tais expressões duram alguns segundos, muitas vezes um único quadro (*frame*). Uma captura instantânea tirada a qualquer momento em que a expressão está em seu ápice pode facilmente transmitir a mensagem da emoção. *É a morfologia, a configuração momentânea produzida pela contração de conjuntos particulares de músculos faciais, que nos informa se estamos diante de raiva, medo, repulsa, tristeza, surpresa ou contentamento.* (EKMAN, 1993, p. 389, meu grifo)

Claro, isto não quer dizer que variações inter-culturais de expressão emocional não possam ocorrer. Em um estudo conduzido por Ekman & Fri-

---

<sup>8</sup> Essas expressões foram cuidadosamente executadas por atores profissionais, baseadas em uma pesquisa metódica sobre os movimentos faciais envolvidos em expressões prototípicas (Ekman & Friesen 1976). As fotografias resultantes fazem parte de um conjunto de estímulos padrão denominado “*pictures of facial affect*”, utilizado até hoje por pesquisadores de todo o mundo.

esen (1975), estudantes japoneses e americanos produziram expressões de emoções negativas ao serem expostos a um filme perturbador enquanto estavam sozinhos em uma sala. No entanto, quando os experimentadores entraram na sala, os americanos continuaram a produzir as mesmas expressões negativas, mas os japoneses passaram a produzir sorrisos educados e outras expressões de emoções positivas. Isto pode ser explicado pelo fato de que há no Japão um tabu contra a expressão de emoções negativas em público (*ibid.*, pp. 23-24). Ekman capturou esses efeitos contextuais com a noção de “regras de exibição”, como o conjunto de hábitos, normas e convenções de cada cultura em relação ao gerenciamento do comportamento emocional em diferentes contextos (*ibid.*, pp. 137-140). Um efeito complementar pode ser capturado com a noção de “regras de decodificação”, como as regras e expectativas culturais em relação à interpretação de fenômenos emocionais (BUCK, 1994).

Desta forma, informação contextual sob a forma de regras de exibição e decodificação pode modular uma atribuição de emoção, mas apenas após a expressão ser percebida automaticamente como uma expressão de uma certa categoria básica. Se estivéssemos presentes no estudo de Ekman e Friesen, teríamos automaticamente reconhecido as expressões dos estudantes japoneses como expressões de contentamento, mas nosso conhecimento das regras de exibição e decodificação naquele contexto nos fariam hesitar em atribuir esta emoção. Embora na teoria das emoções básicas expressões faciais sejam sempre categorizadas em um contexto – dado pela descrição que acompanha as fotografias no paradigma experimental descrito acima –, uma vez que o contexto sugere uma certa emoção básica, a expressão escolhida pelos participantes é invariavelmente uma expressão prototípica desta emoção, não havendo espaço para efeitos contextuais para além das regras de exibição e decodificação.

No entanto, a evidência para a *readout hypothesis* da teoria das emoções básicas seria mais tarde questionada. Um dos primeiros desafios significativos veio da abordagem dimensional de Russell (1997), que produziu evidências claras de efeitos contextuais no reconhecimento de emoções prototípicas básicas. Esta evidência será examinada na próxima seção.

### III. A teoria dimensional

Em uma teoria dimensional, expressões faciais nos fornecem de fato informações sociais valiosas de forma invariável e independente do contexto. Mas ao invés de categorias de emoções básicas serem lidas diretamente da face, há aqui dois tipos de informação que podemos obter invariavelmente de uma expressão facial: por um lado, obtemos informação de que a pessoa está sorrindo, franzindo a testa, arregalando os olhos, etc. Por outro lado, obtemos informação sobre dimensões de valência e excitação, isto é, se a pessoa percebida se sente bem ou mal, e se está relaxada ou excitada. Esses tipos de informação podem ser percebidos na face de forma imediata e automática, independente de fatores contextuais. Com base nestas percepções, usamos informações colaterais disponíveis em um contexto para categorizar a expressão percebida como uma expressão de determinada categoria emocional (Russell 1997).

Nesta teoria, portanto, devemos distinguir entre percepção de informação social e atribuição de emoções. O que percebemos em uma expressão facial são informações físicas e dimensionais, sobre valência e excitação. Usamos então essas informações para inferir, em um contexto, a categoria emocional que melhor se aplica àquela expressão. O contexto intervém apenas no segundo estágio, que é mais lento, demanda mais esforço, e é sensível a diversos tipos de informação colateral, ao contrário do primeiro estágio, que é automático e independente do contexto (CARROLL & RUSSELL, 1996, p. 206).

Para mostrar como a percepção de informação dimensional nos permite categorizar expressões faciais em diferentes contextos, Russell desenvolveu um modelo circunflexo onde expressões faciais estão sistematicamente relacionadas umas às outras por seus graus de valência e excitação (RUSSELL, 1997; ver figura 1 abaixo). Uma categoria emocional particular, como 'raiva', é mapeada em uma região do circunflexo especificada em termos de valência negativa e alta excitação. No entanto, outras categorias, como 'medo', ou 'repulsa', também são mapeadas em regiões espacialmente adjacentes. Isto significa que valores dimensionais de valência e excitação obtidos de uma expressão facial são compatíveis com uma

certa gama de categorias possíveis, mapeadas em regiões adjacentes do circunflexo. Qual categoria particular será aplicada à expressão depende do conhecimento situacional disponível em um contexto de interpretação.

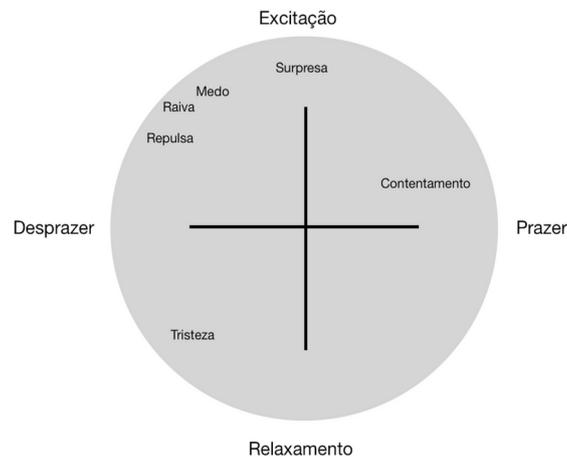


Figura 1

Uma previsão desta teoria é que ao variar a informação situacional disponível, devemos ser capazes de deslocar a categoria atribuída a uma determinada expressão facial, desde que as categorias relevantes pertençam a regiões adjacentes do circunflexo. Para ilustrar com um exemplo, suponha que sejamos apresentados com o seguinte cenário: uma mulher fez reservas para um restaurante disputado com meses de antecedência, mas mesmo assim o *maître* a deixa esperando durante horas, sob o pretexto de que o restaurante está cheio; algum tempo depois, uma celebridade local chega ao restaurante sem ter feito uma reserva, e o *maître* a deixa entrar imediatamente (CARROLL & RUSSELL, 1996, p. 208). Em seguida, somos apresentados a uma fotografia de uma mulher com uma expressão prototípica de medo, e devemos escolher uma categoria de emoção básica que melhor descreva o que a mulher da fotografia deve estar se sentindo no contexto da história. O que irá acontecer?

Se seguirmos a teoria das emoções básicas, a expressão deverá ser automaticamente reconhecida como uma expressão de medo, independente da informação situacional. Embora a resposta de medo seja incongruente com a história, talvez existam informações colaterais que desconhecemos, e que explicam a expressão de medo naquele contexto. No entanto, em uma

teoria dimensional iremos perceber que a mulher da fotografia se encontra em um estado desagradável de alta excitação. Com base nesta percepção, usaremos a informação situacional provida pela história para inferir que a mulher está sentindo raiva.

Para testar ambas as hipóteses, Carroll e Russell (1996) usaram um design experimental onde os participantes eram apresentados com histórias orais de estrutura similar à do restaurante acima, juntamente com fotografias de expressões faciais prototípicas e uma lista de emoções básicas.<sup>9</sup> A tarefa consistia em escolher da lista a palavra que melhor capturava o que a pessoa representada pela fotografia sentia no contexto da história. Como previsto pela teoria dimensional, a informação situacional deslocou a categoria da expressão original para a categoria sugerida pelo contexto, desde que ambas as categorias pertencessem a regiões próximas do circunflexo (CARROLL & RUSSELL, 1996, pp. 212-13). No entanto, quando a tarefa era identificar informações sobre valência e excitação nas expressões faciais, havia um alto nível de consenso independente da variação contextual (RUSSELL, 1997, p. 304).

Estas evidências corroboram a hipótese de que informações sobre valência e excitação podem ser imediatamente percebidas de forma invariável, enquanto a atribuição de categorias de emoções é um processo inferencial sensível ao contexto. Desta forma, enquanto Ekman e Izard julgavam encontrar evidências de universalidade no reconhecimento de emoções básicas, Russell (1997) argumenta que esses experimentos na verdade demonstram universalidade no reconhecimento de valência e excitação; é esta a informação que os participantes de diversas culturas utilizam para combinar uma expressão com uma dada história.

Outros estudos empíricos, no entanto - que serão vistos na seção IV - produziram evidências de que mesmo a percepção de valência e excitação pode ser influenciada por fatores contextuais, e que o contexto pode deslocar a categorização de expressões faciais até mesmo para regiões distantes do circunflexo. Esses dois fatos não podem ser explicados por uma teoria dimensional, como será mostrado na próxima seção.

---

<sup>9</sup> As expressões utilizadas faziam parte do conjunto padrão de Ekman & Friesen (1976).

#### IV. O modelo da percepção de objetos

Uma previsão da teoria dimensional é que o contexto pode deslocar a categorização de expressões faciais somente entre regiões adjacentes do circunflexo. Isto é porque categorias de emoções são mapeadas em regiões definidas por valores dimensionais de valência e excitação, que são compatíveis com mais de uma categoria. Nesse modelo, o contexto não pode fazer com que uma expressão de medo, por exemplo, seja categorizada como uma expressão de tristeza, já que estas categorias são mapeadas em regiões opostas do circunflexo ao longo da dimensão de excitação (ver figura 1 acima).

Mas não é isso o que os estudos de Aviezer e Hassin (2008) encontraram. Nesse design experimental, expressões prototípicas de emoções básicas foram combinadas com posturas corporais inseridas em contextos visuais diversos, que sugeriam categorias de emoções específicas.<sup>10</sup> Uma expressão prototípica de raiva, por exemplo, foi combinada com o corpo de um atleta correndo uma maratona, uma expressão prototípica de repulsa foi colocada em um corpo exibindo punhos cerrados de raiva, e assim por diante. Os participantes foram então apresentados com esses estímulos, e deveriam escolher, a partir de uma lista pré-definida, a categoria de emoção que melhor descrevesse a expressão facial percebida. Esta lista incluía não apenas emoções básicas, mas também outras emoções sugeridas pelo contexto visual, como ‘determinação’ no contexto da maratona. A hipótese dos autores é que o contexto visual extra-facial deveria deslocar a categoria original da expressão para a categoria sugerida pelo contexto, mesmo entre regiões distantes do circunflexo.

Os resultados confirmaram esta hipótese. Uma expressão prototípica de raiva no contexto da maratona foi sistematicamente categorizada como ‘determinação’, uma expressão de repulsa em um contexto de confronto foi sistematicamente categorizada como ‘raiva’, e assim por diante (AVIEZER & HASSIN, 2008, pp. 268-69). Mais diretamente relevante para a presente discussão, uma expressão de tristeza em um contexto de medo foi sistematicamente categorizada como ‘medo’, e avaliada como expressando um alto nível de excitação (AVIEZER & HASSIN, 2008, pp. 276-77). Mas se o con-

---

<sup>10</sup> As expressões utilizadas faziam parte do conjunto padrão de Ekman & Friesen (1976).

texto pode deslocar julgamentos categóricos apenas entre regiões vizinhas do circunflexo, e se informação sobre excitação é lida da face de forma direta e invariável, isso não deveria acontecer.

Em um outro experimento, Aviezer et al. (2012) usaram fotografias de reações expressivas intensas de valência positiva e negativa, como alegria, prazer e vitória do lado positivo, e luto, dor e derrota do lado negativo. Quando os rostos eram vistos isoladamente, as discriminações de valência positiva e negativa não eram melhores do que o acaso. No entanto, quando um rosto que originalmente expressava uma reação positiva era inserida no contexto visual de uma reação negativa, a discriminação de valência era deslocada de acordo com o contexto, e a expressão facial era sistematicamente percebida como uma reação negativa (AVIEZER et al., 2012, p. 1226). Esses estudos mostram que mesmo a percepção de valência e excitação pode ser afetada pelo contexto visual extra-facial, e que portanto estas informações não são lidas do rosto de forma direta e invariável.

Para explicar esses resultados, Aviezer e Hassin cunharam o termo “sementes de emoção” (*emotion seeds*), que se refere a propriedades físicas de uma expressão facial de determinada categoria compartilhadas com expressões de outras categorias (Aviezer & Hassin, 2008, pp. 270-72). Sobrancelhas franzidas, por exemplo, são sementes de emoção de ‘raiva’ compartilhadas com ‘repulsa’. Estas semelhanças físicas podem ser estimadas a partir de modelos computacionais que extraem e calculam similaridades físicas de baixo nível entre expressões prototípicas de forma puramente *bottom-up*, sem conhecimento prévio ou expectativas sobre categorias de emoções (SUSSKIND et al., 2007).

Quando uma expressão facial é vista isoladamente, estas propriedades físicas compartilhadas têm pouca influência sobre a percepção de emoções. No entanto, quando inseridas em um contexto particular, estas propriedades tornam-se ativas, e influenciam a percepção da expressão naquele contexto. Isso explica porque uma expressão prototípica de repulsa pode ser categorizada como uma expressão de raiva, quando inserida em um contexto que sugere raiva: expressões de repulsa e raiva compartilham propriedades físicas (sementes de emoção), que tornam-se ativas quando a expressão é inserida naquele contexto. Nessa teoria, a magnitude e o limite dos

efeitos contextuais podem ser previstos pelos graus de similaridades físicas entre a expressão original e a expressão alvo sugerida em um contexto.

Mas como exatamente o contexto influencia a percepção de emoções em expressões faciais, para além da metáfora de “sementes de emoções que se ativam”? Para responder esta pergunta, Aviezer e Hassin tomam emprestado de Bar (2004) um modelo do viés atencional, onde uma análise rápida do contexto visual em que um objeto é percebido modula processos de reconhecimento do objeto, enviesando a atenção do observador. Aplicado à percepção de emoções, quando a expressão é percebida em um contexto visual, a categoria de emoção sugerida pelo contexto ativa um modelo (*template*) desta categoria. Este modelo irá enviesar a atenção do observador para determinadas micro-regiões da face onde certa atividade muscular é esperada, de acordo com a categoria. Se a atividade muscular for de fato detectada ali, as “sementes” tornam-se ativas, guiando a atenção ao priorizar estas e não outras propriedades físicas na integração perceptiva com outros estímulos visuais. Mas se nenhuma atividade muscular for detectada ali, isto significa que a expressão sugerida pelo contexto não compartilha sementes de emoção com a expressão original, e a percepção será então guiada predominantemente pela categoria da expressão original (AVIEZER & HASSIN, 2008, p. 278).

Evidências empíricas para essa explicação podem ser encontradas em estudos de escaneamento ocular, que mostram que os olhos se movem rapidamente para micro-regiões distintas da face de acordo com a categoria esperada: na repulsa os olhos se movem predominantemente para a região da boca, na raiva eles se movem predominantemente para a região dos olhos, e assim por diante (CALDER et al., 2000a). Se este modelo estiver correto, devemos esperar que variações contextuais induzam alterações em movimentos oculares de acordo com a expressão sugerida pelo contexto, mesmo que a expressão original permaneça constante. De fato, resultados experimentais obtidos a partir de escaneamento visual confirmaram esta previsão (AVIEZER & HASSIN, 2008, p. 280).

Para concluir a seção, neste modelo o papel do contexto na percepção de emoções pode ser compreendido em termos de um processo *top-down* de viés atencional, na medida em que informações visuais extra-faci-

ais (posturas corporais e cena visual) ativam um certo *template* de uma categoria emocional, que enviesará a atenção do agente privilegiando certas configurações faciais em detrimento de outras. Esse modelo assimila a percepção de emoções em expressões faciais à percepção de objetos em cenas visuais, na medida em que a tarefa de reconhecer uma determinada configuração facial pode ser entendida como a tarefa de extrair e integrar diferentes estímulos sensoriais de uma cena visual.

Nesse sentido, não nos surpreende que Aviezer & Hassin se apoiem fortemente sobre o trabalho de Bar (2004). Bar está primariamente interessado em descobrir como o contexto afeta o reconhecimento de objetos em cenas visuais naturais, e descobriu que quando objetos são inseridos em seus contextos esperados (i.e., uma torradeira em um balcão de cozinha), o reconhecimento é significativamente mais rápido e acurado. A proposta de Bar é que uma representação esquemática do contexto (um balcão de cozinha) em áreas pré-frontais torna certas representações (como a de uma torradeira) mais acessíveis, mesmo que outras representações de objetos sejam igualmente compatíveis com as propriedades físicas detectadas (um objeto retangular, de cor preta, etc.). Esta representação enviesa a atenção do observador e prioriza certas propriedades físicas em detrimento de outras, modulando processos de reconhecimento de objetos em áreas corticais inferiores em escalas temporais extremamente rápidas (BAR, 2004: 627). Por esta razão, classificarei esse modelo como “modelo da percepção de objetos”.

O problema em modelar a percepção de emoções na percepção de objetos, no entanto, é que o modelo parece ser capaz de capturar apenas o que acontece no laboratório, onde observadores executam tarefas de categorização a partir de fotografias de expressões e posturas prototípicas. Como Gendron e Barrett sugerem, nesse modelo “*um nariz torcido e uma sobran-celha franzida deve ser repulsa, do mesmo modo que um receptáculo cilíndrico com uma alça deve ser uma xícara*” (GENDRON & BARRETT, 2018, p. 101). Fora do laboratório, no entanto, expressões faciais não são estímulos estáticos, e configurações de musculatura facial e posturas corporais são frequentemente sutis e sob constante alteração, à medida que a interação social se desdobra no tempo. Se isto é verdade, precisamos de modelos que possam capturar a natureza dinâmica da percepção de emoções em expres-

sões faciais, algo que o modelo da percepção de objetos parece ser incapaz de fazer. Esta será a tarefa principal da seção V.

## V. O modelo da sincronia conceitual

De acordo com o modelo da sincronia conceitual, existem dois problemas em modelar a percepção de emoções na percepção de objetos: o primeiro deles, como já visto, é que os estímulos empregados carecem de validade ecológica, por não capturarem a natureza dinâmica, interativa e multidimensional da percepção de emoções em interações sociais. Em segundo lugar, ao focar em expressões e posturas prototípicas, esses modelos não levam em consideração a enorme variabilidade da expressão emocional. Isto é, muitas vezes não arregalamos os olhos quando sentimos medo, nem cerramos e exibimos os punhos quando estamos com raiva (BARRETT, 2009). O medo de uma serpente venenosa pode exibir características expressivas, fisiológicas e fenomenológicas bem diferentes do medo de falar em público ou do medo de uma recessão econômica. E no entanto, mesmo com toda essa variação, as pessoas são perfeitamente capazes de distinguir instâncias variáveis de uma mesma categoria emocional em diferentes contextos (BARRETT, LINDQUIST & GENDRON, 2007).

Portanto, um modelo satisfatório da percepção de emoções deve capturar a natureza dinâmica e variável da expressão e percepção emocional em contextos reais de interação social. No modelo da percepção de objetos, regularidades estatísticas de expressões faciais percebidas ao longo do tempo seriam armazenadas sob a forma de modelos (*templates*) de diferentes categorias de emoção. O modelo da sincronia conceitual, por outro lado, começa com a observação de que categorias de emoções não compartilham regularidades estatísticas significativas através de suas diferentes instâncias, o que significa que o mecanismo proposto acima seria incapaz de gerar resultados satisfatórios fora do laboratório. Ao invés disso, o modelo da sincronia conceitual pressupõe uma teoria da construção psicológica de emoções, onde emoções são concebidas como conceitualizações situadas de sensações intero e exteroceptivas (BARRETT, 2009).

Nesta teoria, emoções não são tipos naturais, mas construções psicológicas. São atos de categorização situada que conferem significado a estímulos sensoriais e sensações interoceptivas em um contexto, baseados em experiências passadas com situações semelhantes. Um conceito de emoção como ‘raiva’ é um modo de categorizar determinadas sensações do corpo e do mundo, que faz com que essas sensações sejam experienciadas como indicações de que alguém naquele contexto cometeu uma ofensa ou um ato injusto. Essas mesmas sensações e percepções, quando categorizadas como ‘medo’ em um outro contexto (ou por uma outra pessoa no mesmo contexto), fazem com que sejam experienciadas como indicações de que algo na situação é ameaçador. E assim por diante (BARRETT, 2009, p. 1293).

Conceitos de emoções, além de situados, são também corporificados e fundamentados em sistemas somatossensoriais (Barsalou 1999). Isso significa que estados sensorimotores, perceptivos e afetivos ativados durante encontros com estímulos que causaram emoções no passado, são capturados e armazenados juntos em áreas associativas que integram diferentes modalidades. Posteriormente, quando o agente se encontra em uma situação similar, ou se lembra de uma experiência prévia, o padrão original de ativação é reativado. Esta reativação, no entanto, é apenas parcial, envolvendo o uso on-line e dinâmico de informações multimodais selecionadas, moldadas e adaptadas à situação em questão (WINKIELMAN, NIEDENTHAL & OBERMAN, 2008).

Um conceito de emoção corporificado, portanto, pode ser compreendido como uma coleção de informações multimodais experienciadas através de instâncias variáveis de uma mesma categoria emocional. Esses conceitos são situados porque um determinado conceito, como ‘medo’, não é ativado isoladamente, mas sempre em uma situação imbuída de significado, que determina quais aspectos de estados sensorimotores e afetivos serão reativados, de acordo com o modo como a atenção está focada e qual informação é relevante para o agente naquele contexto. Quando um conceito de ‘medo’ é ativado durante um encontro com uma serpente venenosa, por exemplo, conceitos como ‘veneno’, ‘morte’, ‘serpente’, etc., serão ativados e integrados de forma significativa ao conceito de ‘medo’, formando o conceito situado ‘medo-de-serpentes’. Alternativamente, se o conceito de ‘medo’ for

ativado durante uma palestra, conceitos como ‘público’, ‘embaraço’, ‘apresentação’, etc., serão ativados e integrados, formando o conceito situado ‘medo-de-falar-em-público’ (WILSON-MENDENHALL et al., 2011, pp. 1107-108).

Cada uma destas conceitualizações leva a diferentes experiências de medo, com características expressivas, fisiológicas e fenomenológicas distintas, que ativam diferentes áreas cerebrais específicas ao processamento da informação mais relevante em cada situação. Estudos de imagens de ressonância magnética funcional mostram que uma experiência de ‘medo’ relacionada à ameaça de dano corporal ativa áreas ligadas ao planejamento de ações no ambiente visuoespacial (córtices parietal inferior, para-hipocampal e cíngulo médio), interocepções (ínsula) e processamento auditivo (giro temporal superior) específicos à ameaça de dano físico. Por outro lado, ‘medo’ em uma situação social como falar em público, ativa preferencialmente áreas ligadas à cognição social, auto-avaliação e controle cognitivo, como os córtices prefrontal ventromedial, orbitofrontal superior esquerdo, e occipital superior (WILSON-MENDENHALL et al., 2011, p. 1124). Conceitos de emoções, portanto, não podem ser compreendidos fora das situações específicas em que são construídos, e envolvem a reativação parcial de estados afetivos e somatossensoriais relevantes àquela situação em particular. É isso o que significa que esses conceitos sejam situados e corporificados.

É importante notar que esses atos de conceitualização situada não são categorizações conscientes e deliberadas, mas tomam a forma de predições geradas pelo modelo generativo do mundo construído pelo cérebro. Como as autoras colocam, “*construir o conceito de ‘repulsa’ é equivalente a prever ‘repulsa’*” (GENDRON & BARRETT, 2018, p. 104).<sup>11</sup> Em uma teoria do processamento preditivo, o cérebro é pensado como uma “máquina de predições”, que está constantemente tentando prever estímulos do mundo e do corpo, a partir de sinais preditivos enviados de níveis corticais mais altos para níveis inferiores seguindo os princípios da probabilidade bayesiana (CLARK, 2013).

---

<sup>11</sup> Na discussão que se segue usarei, portanto, ‘predição’ e ‘conceitualização situada’ de forma intercambiável.

Construir um conceito (situado e corporificado) de ‘medo’ em um contexto é equivalente a gerar uma predição de ‘medo’ específica àquela situação. Estas predições são então corrigidas, ou confirmadas, por sinais *bottom-up* oriundos de áreas somatossensoriais e corticais inferiores (GENDRON & BARRETT, 2018, pp. 102-103). Se, após corrigida, a predição situada de ‘medo’ é aceita como a interpretação mais provável para sensações intero e exteroceptivas naquele contexto, o agente terá uma experiência de medo adaptada e modelada àquele contexto, indicando, por exemplo, perigo iminente de dano corporal ou embaraço social.

Mas embora estejamos até o momento falando sobre a experiência de emoções em primeira pessoa, é importante notar que o mesmo processo ocorre na percepção de emoções em terceiros. Em um contexto de interação social, à medida que o outro está constantemente gerando e corrigindo predições visceromotoras para fazer sentido de suas próprias sensações intero e exteroceptivas, o agente está fazendo a mesma coisa em relação ao outro. Mas, obviamente, as predições do agente não podem ser diretamente corrigidas por interocepções do corpo do outro, mas devem se apoiar em outras fontes de informação. Nesse caso, o agente se apoia em movimentos faciais e corporais, enunciados verbais, e alterações fisiológicas observáveis (sudorese, voz trêmula, etc.), que servem como sinais sensoriais de erro que podem ser usado para ajustar as predições do agente sobre as experiências internas do outro.

Processos de conceitualização situada em experiências subjetivas e percepções de emoções mobilizam portanto os mesmos recursos cognitivos e somatossensoriais; a única diferença está na origem dos sinais sensoriais que constroem e corrigem as predições em cada caso (GENDRON & BARRETT, 2018, p. 104). À medida que a interação social se desdobra no tempo, sinais de erro são propagados e as predições refinadas, ao ponto do agente poder sintonizar suas predições para lidar com toda a variabilidade no comportamento facial e corporal do outro. Este processo é dinâmico e co-construído pelo agente e pelo outro durante a interação social. Além disso, essas predições são corporificadas, gerando ativações de áreas somatossensoriais que buscam simular o estado interno do outro, e corrigidas por

sinais sensoriais a partir do comportamento e alterações fisiológicas observáveis do outro.

Evidências empíricas de que perceber emoções (i.e., construir uma conceitualização situada) envolve simular os estados internos do outro podem ser encontradas em estudos com pacientes com danos a áreas somatossensoriais, que mostram um déficit de performance em tarefas de reconhecimento de emoções (ADOLPHS et al., 2000). Isso é justamente o que esperaríamos se a percepção de emoções fosse baseada na simulação dos estados internos do outro com os recursos somatossensoriais do agente. Além disso, Wicker et al. (2000) mostraram que áreas da ínsula anterior, crucialmente envolvidas na interocepção, são ativadas tanto quando os agentes têm experiências de repulsa, e quando observam repulsa em outros. Os autores tomam esta observação como evidência para a simulação na percepção de emoções. Danos à ínsula também estão ligados a déficits tanto na experiência quanto no reconhecimento da repulsa, o que mais uma vez corrobora a hipótese da simulação (CALDER et al., 2000b).

É importante notar que no modelo da sincronia conceitual, como seu nome indica, a percepção de emoções não é uma questão de acurácia mas de sincronia. Isto é, ao invés de avaliar o sucesso do agente em construir uma conceitualização situada em termos do grau de precisão com que esta conceitualização corresponde ao estado interno do outro, podemos pensar na percepção de emoções como o processo de se estabelecer um grau suficiente de sincronia em relação ao significado de sensações intero e exteroceptivas, através de conceitos situados e corporificados de emoções em um contexto. Como as autoras afirmam,

Ao invés de estudar o ‘reconhecimento’ ou a ‘decodificação’ de sinais não-verbais específicos, sugerimos que a percepção de emoções é melhor estudada enquanto o processo de como dois (ou mais) cérebros chegam a um acordo sobre o significado de sinais sensoriais através de conceitos de emoções (GENDRON & BARRETT, 2018, p. 105).

Isso parece fazer sentido. Afinal, se conceitualizações são situadas e baseadas nas experiências prévias de cada um dos membros da interação, seria implausível requerer uma correspondência entre ambas as conceitualizações. Ao invés disso, a noção relevante aqui é a de *sincronia*, onde um

acordo entre ambas as partes se manifesta na própria dinâmica da interação à medida que previsões são corrigidas e refinadas. Para ilustrar com um exemplo, uma indicação de que duas pessoas convergiram em uma conceitualização situada de 'tristeza', em um contexto, é que o cérebro do agente gerou previsões sensorimotoras para iniciar um toque suave acompanhado de expressões (verbais e não-verbais) de empatia, e que o outro aceitou estas ações.

Para concluir, o modelo da sincronia conceitual admite um papel muito mais radical para o contexto na percepção de emoções do que os outros modelos vistos anteriormente. Aqui, o contexto, sob a forma de dinâmicas de interação social, desempenha um papel constitutivo no processo de conceitualização situada. Isto significa que não seria correto falar em *percepção* de emoções, como se uma expressão facial de emoção fosse um objeto pré-definido por suas propriedades físicas, e o problema fosse apenas explicar como sistemas perceptivos integram estas propriedades em compostos significativos. Ao invés disso, fala-se em co-construção psicológica de emoções em contextos particulares, como um processo de sincronização entre participantes de uma interação social, que se dá através da geração, correção e refinamento constantes de previsões situadas à medida que a interação se desdobra no tempo.

A conceitualização situada particular em que uma díade de interação convergirá é uma função deste processo de sincronização em um contexto, e não pode ser compreendida fora dele. Alguns contextos, por exemplo (como uma relação abusiva ou desgastada), tornam certas categorias, como 'raiva', mais acessíveis do que outras igualmente possíveis. Portanto, o fato de que uma díade sincroniza suas conceitualizações em 'raiva' como uma forma de fazer sentido de sensações intero e exteroceptivas em um dado contexto, não pode ser explicado sem fazer menção à natureza e à dinâmica daquela relação em particular (GENDRON & BARRETT, 2018, p. 106).

Se isto é verdade, então o modelo da percepção de objetos deve ser abandonado. O contexto na percepção de emoções em expressões faciais não é apenas um fator que envia a atenção para uma região ou outra da face; é um elemento constitutivo de como o estado emocional do outro é construído em primeiro lugar. A unidade básica de análise, portanto, não

pode ser o agente solitário e seu sistema cognitivo, mas deve ser a díade (ou o grupo) de interação engajada na atividade situada de co-construção de emoções.

#### **VI. Além da sincronia conceitual: percepção de emoções como engajamentos situados**

Vimos na última seção que para ter validade ecológica, um modelo de percepção de emoções deve capturar a natureza dinâmica e variável de transações emocionais no mundo real. O modelo da sincronia conceitual foi o único examinado aqui que atendeu esse critério, onde a percepção de emoções foi reformulada em termos de um processo de sincronização dinâmica entre conceitualizações situadas de dois (ou mais) participantes à medida que a interação social se desdobra no tempo.

Mas a idéia de olhar para a díade de interação como a unidade básica de análise de fenômenos emocionais não é uma novidade do modelo da sincronia conceitual. Esta ideia já aparece no conceito de *participatory sense-making* desenvolvido por De Jaegher & Di Paolo (2007), além da cognição social em segunda pessoa desenvolvida por Schilbach et al. (2013), onde a percepção de informação socialmente relevante só pode ser compreendida dentro de engajamentos emocionais reais entre pessoas em interações sociais dinâmicas. Schilbach e colaboradores contrastam essa abordagem com o que chamam de uma “visão do espectador” da cognição social, onde um observador desinteressado procura ler o estado mental do outro, que por sua vez não é afetado por esse processo e não pode reagir a ele (SCHILBACH et al., 2013, pp. 394-95).

A abordagem em segunda pessoa, ao contrário, mantém que a cognição social é fundamentalmente diferente quando estamos emocionalmente engajados com alguém, ao invés de apenas observando de fora (este seria o caso em que categorizamos fotografias de expressões faciais em laboratórios de psicologia, por exemplo). Evidências empíricas mostram que as pessoas tendem a classificar de forma significativamente diferente expressões faciais socialmente relevantes quando interagem ativamente com elas, ao invés de apenas observá-las. Engajamentos reais com estímulos sociais também exi-

bem padrões diferenciais de atividade neural, recrutando áreas envolvidas no processamento emocional e avaliativo, em oposição à áreas de processamento visuoespacial ativadas na observação desinteressada (SCHILBACH et al., 2013, p. 400).

O que essas observações sugerem é que a construção de uma categoria emocional em um contexto não é apenas uma reação a um estímulo externo, mas um processo dinâmico de engajamento emocional embebido em uma interação social. A natureza da relação (uma relação abusiva, hierárquica, etc.) e os objetivos da interação social (disputa, insulto, afiliação, etc.), são fatores contextuais que influenciam como as categorias emocionais serão co-construídas, e quais conceitualizações situadas estarão mais acessíveis. Há evidências, por exemplo, de que a raiva é caracterizada por ataque e confronto quando o outro possui um status inferior, mas por evasão e expressão indireta quando o outro possui um status superior; ou que o medo pode nos fazer fugir de um indivíduo ameaçador, mas não quando este também é percebido como amedrontado; e assim por diante (FISCHER & VAN KLEEF, 2010).

Portanto, o modo como percebemos e interpretamos os movimentos expressivos do outro não pode ser previsto apenas por fatos sobre as propriedades físicas do estímulo e sobre a alocação da atenção do agente; deve, ao contrário, fazer um apelo inevitável a fatos sobre a dinâmica da interação social em um contexto particular. O fato de parceiros de longa data sincronizarem rapidamente certas conceitualizações situadas, e não outras, é um resultado direto de longos anos de ajustes mútuos em engajamentos emocionais frequentes, e não pode ser adequadamente explicado fora da natureza desta relação.

Isso nos faz pensar em emoções como sistemas dinâmicos que emergem das interações e relações em que ocorrem. Como dizem Mesquita e Boiger, “*emoções e interações sociais formam um sistema único cujas partes não podem ser separadas*” (2014, p. 298). Mesquita chega ao ponto de propor um revisionismo radical ao estudo das emoções, ao sugerir que o objeto de estudo da área não deve mais ser a *emoção* - que é uma entidade estática - mas sim o *emoting* - que é uma ação situada (2010). Uma ação de *emoting* não é um ato de um indivíduo mas uma propriedade de uma relação, como a

série de “*ajustes contínuos em sintonia com o desdobramento de transações com o mundo prático e social*” (MESQUITA, 2010, p. 89).

Nesta proposta, ao invés de perguntarmos sobre a natureza da emoção, perguntamos sobre as dinâmicas de ações de *emoting*. As predições situadas de um dos membros da díade afetam as predições situadas do outro, criando ciclos de retroalimentação entre os dois que não podem ser reduzidos às emoções individuais de cada pessoa. Podemos então concluir que o contexto aqui desempenha um papel constitutivo na percepção de emoções: simplesmente não há como compreender a percepção de informações socialmente relevantes, como estados emocionais, fora da dinâmica, natureza e objetivos da interação social.

## VII. Conclusões finais

Percorremos um longo caminho desde as primeiras evidências a favor da teoria das emoções básicas, quando a universalidade da expressão e percepção de emoções foi estabelecida simplesmente apresentando às pessoas de diferentes culturas fotografias de atores fazendo expressões prototípicas exageradas (seção II). Uma reação a esta teoria veio da abordagem dimensional de Russell, desenvolvida para lidar com efeitos contextuais que a teoria das emoções básicas não era capaz de acomodar. Nesse modelo, a categoria de emoção atribuída a alguém com base em sua expressão facial depende de conhecimento situacional, embora expressões faciais transmitam informação sobre valência e excitação de forma direta e invariável (seção III). O modelo da percepção de objetos levou o papel do contexto um passo adiante, mostrando que até mesmo a percepção de valência e excitação poderia ser influenciada pelo contexto, sob a forma de posturas corporais e a cena visual na qual a expressão é percebida (seção IV). Para acomodar estas evidências, esse modelo traçou um paralelo com o modo como cenas naturais influenciam a percepção e o reconhecimento de objetos, ao ativar um modelo (*template*) que envia a atenção do observador para determinadas propriedades físicas em detrimento de outras.

Esse modelo, no entanto, não sobrevive a um confronto com interações sociais reais fora do laboratório, onde expressões faciais e postu-

ras corporais não são estímulos estáticos mas movimentos dinâmicos, sutis e variáveis, que estão em constante transformação à medida que a interação social se desdobra no tempo. A percepção de emoções não é uma tarefa puramente perceptiva, mas uma tarefa de cognição social situada, na medida em que fenômenos emocionais emergem das dinâmicas da interação social entre duas (ou mais) pessoas. Como argumentado por autores como De Jaeger & Di Paolo (2007) e Schilbach et al. (2013), a cognição social é fundamentalmente diferente quando duas pessoas estão ativamente engajadas, ao invés de meramente observando estímulos sociais de uma perspectiva desinteressada, algo que recruta recursos cognitivos diferentes e gera resultados experimentais diferentes.

Se isto é verdade, precisamos de modelos de percepção de emoções que estejam aptos a capturar a natureza dinâmica, social e variável de fenômenos emocionais. Na seção V introduzi o modelo da sincronia conceitual desenvolvido por Gendron e Barrett (2018), dentro de uma teoria da construção psicológica de emoções. Neste modelo, a percepção de emoções não é um processo separado que ocorre após o estabelecimento de um estado emocional, mas ambos são vistos como parte de um mesmo processo de co-construção, à medida que duas (ou mais) pessoas engajadas em uma interação social sincronizam suas conceitualizações situadas para coordenar seus comportamentos verbais e não-verbais. Esta sincronia é obtida pela geração constante de predições por cérebros individuais, mas que são corrigidas e refinadas por sinais de erro provenientes do outro, até se chegar a uma sincronia conceitual mutuamente aceita. Como predições e sinais de erro emergem no contexto de uma interação, e são moldados e determinados pelas dinâmicas da relação, esse é um argumento convincente para considerar o contexto não apenas como uma força modulatória ou facilitadora da percepção de emoções, mas como desempenhando um papel constitutivo nesta.

A esse respeito, concentrei-me no modelo da sincronia conceitual porque ele oferece uma explicação mais detalhada do que é a conceitualização situada, ao se apoiar sobre uma teoria bem estabelecida e empiricamente informada sobre conceitos situados e corporificados (BARSALOU, 1999), que seriam construídos através de predições e correções seguindo os princípios da probabilidade bayesiana em uma teoria do processamento preditivo,

também já bem aceita e desenvolvida (CLARK, 2013). Mas outros modelos brevemente descritos na seção VI, tal como o modelo sociodinâmico de Batja Mesquita (MESQUITA 2010; MESQUITA & BOIGER, 2014), também possuem os recursos adequados para capturar esses fatos de forma satisfatória.

Esse programa de pesquisa aponta para uma nova forma de se pensar em fenômenos emocionais, onde a explicação de como percebemos e atribuímos significado a expressões faciais e gestos corporais se dá apenas dentro de um contexto de interação social. Em outras palavras, categorias de emoções não são encontradas dentro de nós, mas sim entre nós, na dinâmica de nossas interações.

### Referências

ADOLPHS, R, et al. A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *Journal of Neuroscience*, v. 20, pp. 2683–2690. 2000.

AVIEZER, H., & HASSIN, R.. Putting facial expressions back in context. In: Nalini Ambady, John Skowronsky (eds.). *First Impressions*. New York, London: The Guilford Press, pp. 255-286. 2008.

AVIEZER, H., et al. Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, v. 338, n. 6111, pp. 1225-1229. 2012.

BAR, M. Visual Objects in Context. *Nature Reviews Neuroscience*, v. 5, n. 8, pp. 617-629. 2004.

BARRETT, L.F. Variety is the spice of life: A psychological construction approach to understanding variability in emotion. *Cognition & Emotion*, v. 23, n. 7, pp. 1284-1306. 2009.

BARRETT, L.F. Was Darwin wrong about emotional expressions? *Current Directions in Psychological Science*, v. 20, n. 6, pp. 400-406. 2011.

BARRETT, L. F., LINDQUIST, K., & GENDRON, M. Language as a context for emotion perception. *Trends in Cognitive Sciences*, v. 11, pp. 327-332. 2007.

BARSALOU, L. W. Perceptual symbol systems. *Behavioral & Brain Sciences*, v. 22, pp. 577-660. 1999.

BUCK, R. Social and emotional functions in facial expression and communication: the readout hypothesis. *Biological Psychology*, v. 38, n. 2-3, pp. 95-115. 1994.

CALDER, A.J., et al. Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, v. 26, pp. 527-551. 2000a.

CALDER, A.J., et al. Facial expression recognition by people with Mobius syndrome. *Cognitive Neuropsychology*, v. 17, pp. 73-87. 2000b.

CARROLL, J.M., & RUSSELL, J.A. Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology*, v. 70, pp. 205-218. 1996.

CLARK, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, v. 36, pp. 281-253. 2013.

CRIVELLI, C., & FRIDLUND, A. Inside-Out: From Basic Emotions Theory to the Behavioral Ecology View. *Journal of Nonverbal Behavior*, v. 43, n. 2, pp. 161-194. 2019.

DARWIN, C. *The expression of emotion in man and animals*. New York: Appelton. 2005 (originalmente publicado em 1872).

DE JAEGHER, H., & DI PAOLO, E. Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6, pp.485-507.

EKMAN, P. The argument and evidence about universals in facial expressions of emotion. In: Hugh Wagner & Antony Manstead (eds). *Handbook of Social Psychophysiology*. Hoboken: Wiley-Blackwell, pp. 143-164. 1989.

EKMAN, P. An Argument for Basic Emotions. *Cognition & Emotion*, v. 6, n. 3-4, pp. 169-200. 1992.

EKMAN, P. Facial expression and emotion. *American Psychologist*, v. 48, n. 4, pp. 384-392. 1993.

EKMAN, P., & FRIESEN, W. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, v. 17, n. 2, pp. 124-129. 1971.

EKMAN, P., & FRIESEN, W. *Unmasking the Face*. Englewood Cliffs: Prentice-Hall. 1975.

EKMAN, P., & FRIESEN, W. *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press. 1976.

FARRONI, T., et al. The perception of facial expressions in newborns. *European Journal of Developmental Psychology*, v. 4, n. 1, pp. 2-13. 2007.

FISCHER, A., & VAN KLEEF, G. Where have all the people gone? A plea for including social interaction in emotion research. *Emotion Review*, v. 2, n. 3, pp. 208-211.

GENDRON, M., & BARRETT, L.F. Reconstructing the past: A century of ideas about emotion in psychology. *Emotion Review*, v. 1, n. 4, pp. 316-339. 2009.

GENDRON, M., & BARRETT, L.F. Emotion Perception as Conceptual Synchrony. *Emotion Review*, v. 10, n. 2, pp. 101-110. 2018.

LEDOUX, J. *The Emotional Brain, 2nd edition*. London: Phoenix. 1998.

MESQUITA, B. Emoting: A Contextualized Process. In: Batja Mesquita, Lisa Feldman Barrett & Eliot R. Smith (eds.). *The Mind in Context*. New York, London: The Guilford Press, pp. 83-104. 2010.

MESQUITA, B., & BOIGER, M. Emotions in context: A sociodynamic model of emotions. *Emotion Review*, v. 6, n. 4, pp. 298-302. 2014.

RUSSELL, J. Reading emotions from an into faces : Resurrecting a dimensional-contextual perspective. In: James Russell & José Miguel Fernandez-Dols (eds.) *The Psychology of Facial Expression*. Cambridge: Cambridge University Press, pp. 295-320. 1997.

SCHILBACH, L., et al. Towards a second-person neuroscience. *Behavioral and Brain Sciences*, v. 36, n. 4, pp. 393-414. 2013.

SHARIFF, A., & TRACY, J. What are emotion expressions for? *Current Directions in Psychological Science*, v. 20, n. 6, pp. 395-399.

SUSSKIND, J.M., et al. Human and computer recognition of facial expressions of emotion. *Neuropsychologia*, v. 45, pp. 152-162. 2007.

WICKER, B., et al. Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron*, v. 40, pp. 655-664. 2003.

WILSON-MENDENHALL, C., et al. Grounding emotion in situated conceptualization. *Neuropsychologia*, v. 49, n. 5, pp. 1105-1127. 2011.

WINKIELMAN, P. NIEDENTHAL, P., & OBERMAN, L. The Embodied Emotional Mind. In: G.R. Semin & E.R. Smith (eds.). *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches*. New York: Cambridge University Press, pp. 263-288. 2008.