My main goal will be to introduce the framework of psychological game theory. To prepare, start out reading sections 1 and 2 (including footnotes 2-6) of this article:

Battigalli, Pierpaolo & Martin Dufwenberg (B&D) (2009), "Dynamic Psychological Games," *Journal of Economic Theory* 144, 1-35.

Then read also the rest + following two papers:

Battigalli, P. & M. Dufwenberg (2019), "Psychological Game Theory," IGIER WP 646.

Battigalli, P., Roberto Corrao & M. Dufwenberg (2019), "Modeling Belief-Dependent Motivation in Games," *Journal of Economic Behavior & Organization,* forthcoming.

Carefully study, reflect on, and understand how these articles relate to the earlier work by Geanakoplos, Pearce & Stacchetti (1989, *GEB*).

In my lectures I will first use guilt aversion as my leading example/application. I'll touch on Dufwenberg's (2002, *JEBO*) example of psychological forward induction, Dufwenberg & Gneezy's (2000, *GEB*) early experimental test, Charness & Dufwenberg's (C&D) (2006, *Econometrica*; 2010, *EL*; 2011, *AER*) considerations of communication, B&D's (2007, *AER P&P*) general model, Attanasi, Battigalli & Manzoni's (2016, *Management Science*) extension to incomplete information, Dufwenberg, Gächter & Hennig-Schmidt's (2011, *GEB*) (DG&HS) work on framing, Battigalli, Charness & Dufwenberg (2013, *JEBO*) account of how the model can explain aspects of Uri Gneezy's (2006, *AER*) deception data, Chang, Smith, Dufwenberg & Sanfey's (2011, *Neuron*) neuroeconomics turn, and other papers, including some forthcoming in the special issue on psychological game theory that Amrish Patel are editing for *JEBO*, all of which are cited in B&D (2019, *op. cit.*). Most of the papers where I was a co-author (+also papers mentioned below) are downloadable on my homepage:

http://www.u.arizona.edu/~martind1/

Next, read up on perceived cheating aversion, the topic I'll cover next in my lectures:

Dufwenberg, M. & M.A. Dufwenberg (D&D) (2018), "Lies in Disguise – A Theoretical Analysis of Cheating," *Journal of Economic Theory* 175, 248-64.

Exercises: 1. Prove that $s \in S$ defined by $s(x)=x+1$ for all $x<n$ and $s(n)=n$ is not a SE for any θ. 2. Prove that $s \in S$ defined by $s(x)=n$ for all x is not a SE for any high enough θ. 3. An erroneous claim is made on page 257 – find it!

Note that Gneezy, Kajackaite & Sobel (2018, *AER*) present a related theory. Have a look, compare assumptions, and appreciate key differences (see D&D's section 6 for commentary).

Next step: Learn, test, and apply reciprocity theory! Start out with Dufwenberg & Kirchsteiger (D&K) (2004, *GEB*) – work through all examples carefully! Note and understand how D&K's (2004) article relates to the earlier work by Rabin (1993, *AER*). Then look up also D&K (2000, *EER*), DG&HS, Dufwenberg, Smith & Van Essen (2013, *EI*) (DS&VE), van Damme et al (2014, *JEBO*; section 6), Dufwenberg & Rietzke (2016, mimeo), Dufwenberg & Patel (2017, *GEB*), and Jang, Patel & Dufwenberg (2018, *GEB*). Further references are given in D&K (forthcoming *JEBO*, footnote 1).

<u>Then</u>, study…
- anger theory: read Battigalli, Dufwenberg & Smith (forthcoming, *GEB*) (BD&S);
- social norms & respect: read Dufwenberg & Lundholm (2001, *EJ*) and also Bernheim (1994, *JPE*), and note also the related discussion in B&D (2009) and D&D.
- B&D (2019) again, re-reflecting on additional topics (disappointment, regret, self-esteem…).

<u>Finally, solve more exercises:</u>

4. Consider the Abi-&-Ben example on p5 in B&D (2009), but allow that Abi's utility is w-m-$\theta_A \cdot \max\{\mu\text{-}m,0\}$, where $\theta_A \geq 0$ is Abi's guilt sensitivity. (μ is the same as B&D's μ with an upper-bar; I just can't make upper-bar in Word). B&D consider the case where $\theta_A=2$. What would Abi choose if (i) $\theta_A=3$, $w=100$, and $\mu=5$? (ii) $\theta_A=3$, $w=5$, and $\mu=100$? (iii) $\theta_A=3$, $w=100$, and Abi assigns probability ½ to Ben expecting nothing and 5? (iv) $\theta_A=0$, $w=100$, and $\mu=5$? (v) $\theta_A=0$, $w=5$, and $\mu=100$?

5. Consider the game form in Figure 1 of C&D. Apply B&D's (2007) theory of simple guilt. For which values of θ_B is (*In*, *Roll*) an SE?

6. Now apply B&D's (2007) theory of guilt-from-blame. (a) For which values of θ_B is (*In*, *Roll*) an SE? (b) If $\sigma=(\sigma_A,\sigma_B)$ is an SE, describe $\sigma_B(Roll)$ as a function of θ_B.

7. Consider the following game under D&K's reciprocity theory:

	C	***D***
C	3, 3	0, 4
D	4, 0	1, 1

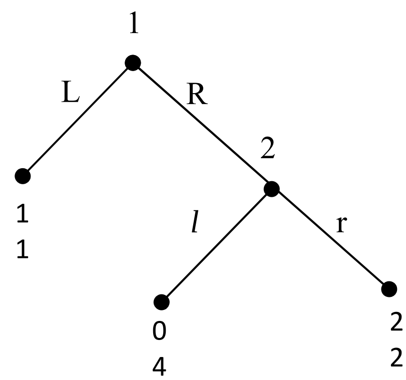
Let $Y=Y_{12}=Y_{21}\geq 0$. (i) For which values of Y is (C,C) an SRE? (ii) What about (D,D)? (iii) (D,C)?

8. Consider instead the following game:

	A	***B***
A	2, 1	0, 0
B	0, 0	1, 2

Again $Y=Y_{12}=Y_{21}\geq 0$. (i) For which Y-values are (A,A) an SRE? (ii) (B,B)? (iii) (B,A)? (iv) (A,B)?

9. Solve the following game for all its SREs for all values of Y_{12} and Y_{21}:

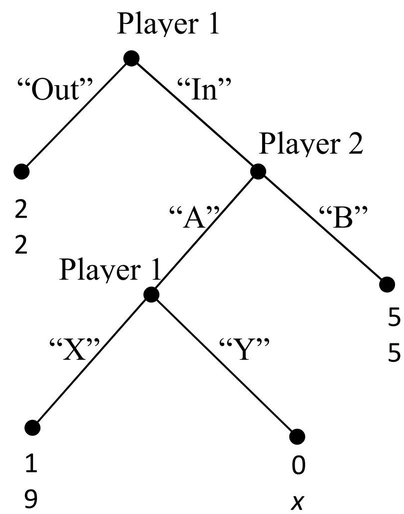


10. Find all Nash equilibria in DG&HS' let's-get-7 and let's-get-9 games (page 460). Then consider the public goods game subsequently explored and explain why, under guilt aversion as well as reciprocity theory, if, for all i, γ_i and Y_i are high enough, everyone choosing 0 as well as everyone choosing 20 is consistent with equilibrium play as defined by B&D and by D&K.

11. Consider the definition of vengeance equilibrium in DS&VE section IV.B. Note that "neither SRE nor VE is a refinement of the other concept," and convince yourself that you understand why by studying DS&VE's example in their Figure 2. Then move instead to another paper, namely van Damme et al (2014) (available on my homepage) and its section 6. Study the two observations there, and then explain why or why not those propositions would change if one applied the VE concept instead of the SRE concept to the ultimatum game.

12. DS&VE discuss a "miserable VE," by which they mean that behavior in some subgame is characterized by the players being unkind to each other despite that there is a different VE where they would be kind. Similar issues arise with SRE; see van Damme et al (2014, section 6, Observation 2(b)). Construct another example based on some other game, as simple as possible, which exhibits miserable play in SRE.

13. Consider the following game. (i) Find all subgame perfect equilibria assuming selfishness. Does x "matter"? Why/why not?



(ii) For which values of x can ((In,Y),B) be supported as a vengeance equilibrium (VE), and for what θ_{ij}-values? (iii) Suppose the game were instead just the subgame where 1 chooses between X and Y; for which values of $x \geq 0$ can Y be supported as a VE?

14. Check out BD&S. (i) Study example 1, in particular the last three lines where some formalism is illustrated. Suppose that the game form in Figure E had been substituted for that in Figure A. How would the stuff in the last three lines of Example 1 have looked instead? (ii) Study Example 10, based on the hammering-1's-thumb game of Figure C. Consider the set $(1/(2(1-\varepsilon)), 1/(2/(1-2\varepsilon))$ mentioned towards the end. Suppose that the game form in Figure C had had payoffs of (1,1) substituted for where it now says (1,2). What set would subsequently have appeared instead of $(1/(2(1-\varepsilon)), 1/(2/(1-2\varepsilon))$, towards the end of the accordingly modified Ex 10?

15. Consider the ultimatum minigame in Figure A of BD&S. (i) According to D&K's theory, for which values of Y_{ab} and Y_{ba} is (f,n) an SRE? (ii) What about (g,y)? (iii) What about (g,n)? Now consider instead BD&S theory of simple anger. (iv) For which values of θ_b is (f,n) an sequential equilibrium (SE)? (v) What about (g,y)? (vi) What about (g,n)? (vii) How do the answers to (iv)-(vi) change if anger-from-blame is considered instead of simple anger?

16. Consider Example 2 of BD&S. Explain why the last sentence must be true. Then consider Example 6 and explain how the value of $\alpha_a(N|B)= 1/(2\varepsilon\theta_a)-(1-2\varepsilon)/\varepsilon$ was obtained.

17. Consider again D&D, assuming that n=5 (the die-throw case). (i) Suppose that $s(x)(5)=1$ for all x. Interpret that strategy. For which values of θ is it a SE? (ii) Suppose that $s(x)(x)=1$ for all x. Interpret that strategy and explain for which θ-values it's a SE. (iii) W.r.t. p. 253, explain why if s is a SE that has "full-support-on-y" then the "0-utility-property" must hold.