

## Dijital Antisosyal Davranışların Kültürel Temelleri: Sosyal Fayda ve Maliyet Eksenli Tercih Fonksiyonları Çerçevesinde Çok Düzeyli Bir Analiz

Ömür Saltık<sup>1</sup>

Filiz Bozagaç<sup>2</sup>

### Özet

Dijital antisosyal davranışlar, nefret söyleminden trollüğe uzanan geniş yelpazede, salt bireysel patolojilerin veya ahlaki birer sapmanın ötesinde, kültürel olarak yapılandırılmış tercih fonksiyonlarının birer çıktısıdır. Bu eylemler, dijital ekosistemin mimarisiyle şekillenen, statü, kimlik inşası ve anlık psikolojik doyum gibi marjinal faydalar ile dışlanma gibi sosyal maliyetler arasındaki dinamik bir hesaplaşmaya dayanmaktadır. Dijital platformların anonimlik ve düşük hesap verebilirlik sunan yapısı, bireysizleşmeyi (kalabalık içinde kimliğini kaybetme) körükleyerek sosyal maliyetleri asgariye indirirken, algoritmik görünürlük negatif dışlıkları aktör için rasyonel birer kazanıma dönüştürmektedir. Ancak bu rasyonalite evrensel olmamakla beraber “Bireycilik, Güç Mesafesi ve Norm Sıklığı gibi kültürel boyutlar tarafından meşruiyet sınırları sistematik olarak yeniden çizmektedir. Örnek olarak, bireyci yapılarda ghosting veya trollük şahsi bir özerklik alanı olarak tolere edilirken, kolektivist toplumlarda aynı eylem ağır sosyal ihlal sayılmakta, ancak kolektif linç, paradoksal bir şekilde toplumsal düzeni koruma aracı olarak kutsanabilmektedir. Sonuç olarak, dijital saldırganlık, kimlik ekonomisi ile teknolojik teşviklerin birleştiği noktada gerekçelendirilen ve toplumsal bağlamda yerleşmiş bir tercih olarak ortaya çıkmaktadır. Bu nedenle etkili bir mücadele, salt yasaklayıcı önlemleri değil, failerin fayda-maliyet dengesini kolektif refah lehine onaracak disiplinlerarası bir politika mimarisini zorunlu kılmaktadır. Bu perspektif, dijital şiddetin kök nedenlerini dijital platformların tasarımı ve kültürel rasyonalite düzleminde yeniden konumlandırmayı amaçlamaktadır.

1 Dr. Öğr. Üyesi, Dokuz Eylül Üniversitesi, omursaltik09@gmail.com , <https://orcid.org/0000-0001-8507-8971>

2 Dr., Bağımsız Araştırmacı, fbozagaç@gmail.com, <https://orcid.org/0000-0002-3764-0111>

## 1. Dijital Antisosyal Davranışların Kavramsal Çerçevesi

Dijitalleşme, toplumsal etkileşim kodlarını sarsarken, antisosyal davranışlar sadece görünürlük kazanmamış, etki alanlarını benzersiz kılmıştır. Bunlar sadistik kişilik dışavurumları (Buckels v.d., 2014) veya durumsal sapmalar olarak yorumlanabilmektedir (Moffitt, 1993; Cheng vd., 2017). Teknolojik imkanlar, sorumluluk bilincini seyrelten bir kimliksizleşme zemini sunmaktadır (Hossain v.d., 2025; Suler, 2004). Algoritmik yangın fırtınaları ise (Wilczek, 2016; Zuboff, 2019), normları koruma adına yıkıcı sürü reflekslerini rasyonelleştirir (Fehr & Gächter, 2002). Bu platformlar artık stratejik “alay edilebilirliğin” müzakere edildiği özgün sahalara ve sürekli bir meşruiyet baskısının üretildiği alanlara dönüşmüşlerdir (Ren vd., 2024; Kaplan, 2021). Süreç mikroiktisadi olarak fayda maksimizasyonu açısından statü ve güç gibi maddi olmayan unsurları da içselleştiren ve kültürel olarak kodlanmış tercih fonksiyonlarına ulaşmaktadır (Akerlof & Kranton, 2000; Becker, 1996). Sosyal Oyun Teorisi ise bu denklemi “yetkinlik” ve “ahlak” algıları ekleyerek karmaşıklaştırır (Krueger v.d., 2020). Sonuçta bu davranışlar, marjinal fayda ve maliyet dengesinde eylemi yapan için rasyonel birer hamle haline dönüşmektedir. Nitekim bu platformlarda troller olarak adlandırılan kullanıcıların “lulz” sadistik hazı içsel bir ödül olarak kabul edilmektedir (Buckels v.d., 2014). Anonimlik, sosyal maliyeti sıfırlarken, gelecek miyopluğu kullanıcının anlık ilgiyi itibar kaybına tercih etmesine neden olur (Hossain, v.d., 2025; Mazas, v.d., 2000). Bu yapı negatif sosyal tercihlerle açıklanmaktadır (Olschewski vd., 2019; Iwasaki, 2023; Levinson & Peng, 2006).

Kültür ise bu noktada fayda ve maliyet analizinin yapıldığı zihinsel muhasebe kuralları ile kullanıcının risk toleransını belirler (Gelfand v.d., 2011; Hofstede, 2001). Örnek olarak, bireyci kültürlerde ghosting düşük maliyetli olarak algılanırken, kolektivist yapılarda bu davranışın ağır bir kimlik kaybı ürettiği kabul edilir (Hofstede, 2001; Akerlof & Kranton, 2000). Altruistik cezalandırma olarak rasyonelleştirilen linç kültürü ise aslında eski refleksler ile dijital hızın evrimsel uyumsuzluğudur (Fehr & Gächter, 2002; Gelfand, vd., 2024). Bu anlamda dijital uygulamalarda fayda ve maliyet dengesi evrensel bir matematikten ziyade kültürel ipuçlarıyla şekillenmektedir (Levinson ve Peng, 2006). Otoriteyi pekiştiren manipülasyonun yüksek güç mesafesinde “işlevsel” görülmesi bu özgün rasyonelitenin bir yansımasıdır (Hofstede, 2001). Suç Üçgeni ekseninde bu eylemler, aldatıcı bir hayatta kalma stratejisi olarak kodlanmaktadır (Ren v.d., 2024). Norm sıklığı yaptırımların öngörülebilirliği üzerinden antisosyal eğilimi sönmümlendirirken (Gelfand vd., 2024), belirsiz sahalarda bireysel yansıtma yoluyla bir meşruiyet yanığı inşa eder (Briscese v.d., 2024). Psikolojik analizler tek başına yetersiz kalırken birey, dijital anonimliğin düşürdüğü sorumluluk maliyetiyle kimliksizleşme sürecine girerek (Hossain

vd., 2025), özellikle iptal kültüründe gördüğümüz itibar kaskadlarının sürüklediği rasyonel “yangın fırtınalarının” birer parçası haline gelmektedir (Traversa, 2022; Wilczek, 2016). Bu eylemler bireysel sapma değil, yapısal teşviklerin ürettiği kolektif pratikler olarak karşımıza çıkmaktadır.

*Tablo 1. Dijital Antisosyal Davranışların Temel İşlevi, Kültürel Bağlamda Baskın Dinamikler ve Tercih Fonksiyonlarının Kavramsal Haritası*

Dijital Antisosyal Davranış	Temel İşlevi	Kültürel Bağlamda Baskın Dinamik	Tercih Fonksiyonu
Nefret Söylemi (Akerlof & Kranton, 2000)	Öteki yaratma, sınır çizme, kimlik pekiştirme	Kolektivizm (Hofstede, 2001), norm sıklığı (Gelfand v.d., 2011), belirsizlikle başetme (Hofstede, 2001) (yüksek tehdit algısı)	Grup aidiyeti ve kimlik pekiştirmeden elde edilen fayda (Gelfand vd., 2024), olası sosyal yaptırım maliyetinden daha yüksek algılanmaktadır.
Ghosting (Ibáñez vd., 2016)	Kaçınma, duygusal yükten kurtulma	Bireycilik (Hofstede, 2001), düşük bağlanma normları (Gelfand v.d., 2011),	Duygusal maliyetin azaltılmasından sağlanan fayda (White vd., 2013), ilişkiyi sürdürmenin sosyal maliyetinden daha yüksek görülmektedir.
Çevrim içi Linç (Fehr & Gächter, 2002)	Norm koruma, kolektif cezalandırma	Kolektivizm (Hofstede, 2001), norm sıklığı (Gelfand v.d., 2011), maskülenlik (Hofstede, 2001)	Kolektif onay ve ahlaki üstünlükten elde edilen fayda (Wilczek, 2016; Gelfand vd., 2024), bireysel sorumluluk ve suçluluk maliyetini gölgede bırakmaktadır
Doxxing (Olschewski vd., 2019)	Güç gösterisi, korkutma, kontrol	Güç mesafesi (Hofstede, 2001), otorite meşruiyeti	Kontrol ve caydırıcılık yoluyla elde edilen güç faydası (Akerlof & Kranton, 2000), hukuki ve etik yaptırım riskinden daha yüksek algılanmaktadır (Olschewski vd., 2019).
İptal Kültürü (Traversa, 2022)	Ahlaki düzenleme, dışlama	Norm sıklığı (Gelfand v.d., 2011), kolektif ahlak (Hofstede, 2001)	Norm ihlalini cezalandırmaktan doğan kolektif fayda (Gelfand v.d., 2011), bireysel zarar ve orantısızlık maliyetlerinden daha yüksek değerlendirilmektedir.
Gaslighting (Seligman, 1975)	Gerçeklik kontrolü, psikolojik üstünlük	Güç mesafesi (Hofstede, 2001), hiyerarşik ilişkiler	Algı kontrolü ve üstünlükten sağlanan fayda (Akerlof & Kranton, 2000), ilişkinin bozulma ve güven kaybı maliyetinin üzerinde görülmektedir.
Siber Zorbalık (Buckels vd., 2014)	Üstünlük kurma, görünürlük	Maskülenlik (Hofstede, 2001), rekabet kültürü	Görünürlük ve statü kazanımının marjinal faydası (Krueger, 2020), mağdura verilen zararın sosyal maliyetinden daha yüksek algılanmaktadır.

Dijital Antisosyal Davranış	Temel İşlevi	Kültürel Bağlamda Baskın Dinamik	Tercih Fonksiyonu
Trollük (Buckels vd., 2014)	Provokasyon, dikkat ve statü	Bireycilik (Hofstede, 2001), norm esnekliği (Gelfand v.d., 2011)	Eğlence, dikkat ve görünürlükten elde edilen fayda (Hossain vd., 2025; Kaplan, 2021; Buckels vd., 2014), yaptırım ve itibar kaybı maliyetinden daha baskın hâle gelmektedir. Davranışları yönlendirme ve sonuçları kontrol etmenin faydası (Olschewski vd., 2019), etik ihlal ve güven erozyonu maliyetinden daha yüksek algılanmaktadır (Schwartz, 2012).
Manipülasyon	Algı yönetimi, yönlendirme	Güç mesafesi (Hofstede, 2001), otoriteye itaat	

## 2. Antisosyal Davranışların Çok Katmanlı Doğası

Dijital antisosyal eylemler, biyolojik eğilimler ile toplumsal yapıların sıkı etkileşimi ( $G \times E$ ) içinde şekillenen çok katmanlı olgulardır (Burt, 2022; Gelfand vd., 2024; Suler, 2004). Anonimlik ve dağınık kolektivite bireyde kimliksizleşme yaratarak, normalde baskılanan dürtülerin maliyetini düşürmektedir (Hossain vd., 2025). Farklı biçimlerde ortaya çıkan bu pratikler, özünde güç kurma ve aidiyet üretme süreçlerinde birleşmektedir (Tajfel vd., 2001). Troller ise “lulz” olarak tanımlanan sadistik bir hazdan yararlanırken, bazı aktörler cezalandırmanın düşük duygusal maliyetini bir fayda kalemi olarak hesaplamaktadır (Buckels vd., 2014; White vd., 2013). Dolayısıyla süreci kavramak, risk algılarının kültürel kodlarla yeniden ölçeklendiği psikolojik, sosyolojik, kültürel ve teknolojik boyutları birlikte ele alan disiplinlerarası ve bütüncül bir perspektifi zorunlu kılmaktadır (Levinson & Peng, 2006; Costello vd., 2019; Hofstede, 2001).

### 2.1. Psikolojik ve Nörobiyolojik Dinamikler

Dijital antisosyal davranışların kökenindeki empati eksikliği, saldırganlığı tercih fonksiyonunda bir “fayda” kalemi olarak kodlamaktadır. Trollerin sadistik hazı (“lulz”) marjinal getiriye artırırken (Buckels vd., 2014), “duygusuz-hissiz” özelliklere sahip bireylerin hissetmediği içsel maliyet, eylemi “ekonomik olarak ucuzlatmaktadır” (White vd., 2013). Sosyal Kimlik Teorisi ise aidiyeti bir kimlik faydası ( $I_j$ ) olarak kurgulayıp dış grup nefretini rasyonelleştirirken (Tajfel vd., 2001; Akerlof & Kranton, 2000). Dijital linç ve iptal kültürü eylemleri, bireylere “kolektif trollük” sunarak saldırganlığı rasyonel bir grup aktivitesine dönüştürmektedir (Hossain vd., 2025).

Psikanalitik açıdan antisosyal davranışlar, bastırılmış korku ve öfkenin “öteki”ne yansıtılmasıyla meşrulaşmaktadır (Freud, 2015). Öfke duygusu, sorumluluğu durumsal etkenlerden ziyade ısrarla bireyde aramayı şiddetlendirir (Keltner vd., 1993). Yüksek asabiyet ise gelecek miyopluğuna yol açarak ceza riskinin zihinlerde canlandırılmasını engeller (Blair vd., 2020). Dijital platformlar, bu kontrolsüz duyguların düşük maliyetle dışa vurulabildiği mecralara dönüşmektedir.

Bağlanma kuramı ise ghosting gibi pratikleri mikroiktisadi bir mantığa oturtmaktadır. Kaçınmacı bireyler, ilişki maliyetten sakınmak adına iletişimi ani şekilde sonlandırabilmektedir. Bu durum, oyun teorisindeki “Parayı Al ve Kaç” stratejisiyle örtüşen tek taraflı bir sözleşme ihlalidir (Ibáñez vd., 2016; Bowlby vd., 1992; Shaver & Hazan, 1987; Balafoutas vd., 2021). Gaslighting süreci ise karşılık vermeme eğilimiyle uyumlu olarak, bireylerin gerçeklik algısını sistematik biçimde zayıflatır ve öz yeterlik kaybını besler (Sweet, 2019; Christie & Geis, 2013; Seligman, 1975).

Nörobiyolojik açıdan dijital anonimlik, duygu-biliş etkileşimi modeline paralel olarak, amigdala üzerindeki prefrontal kontrolü zayıflatarak saldırganlığın önündeki bilişsel engelleri kaldırabilmektedir. Örnek olarak, prefrontal korteks (vmPFC) işlevlerinde bozukluk veya lezyon olan antisosyal aktörlerde ‘duygusal entegrasyon’ eksik olduğu için, başkasına zarar vermenin vicdani maliyet hesapları yapılamaz (Pessoa, 2008). Örnek olarak, prefrontal korteks (vmPFC)’de lezyonu olan antisosyal aktörler, “duygusal entegrasyon” eksikliği nedeniyle başkasına zarar vermenin vicdani maliyet hesapları eksiktir. Bu durum, yıkıcı kararları nöral düzeyde ekonomik olarak ucuzlatmaktadır (Koenigs vd., 2010; White vd., 2013). Kortizol ve adrenalin saldırganlığı artırırken, testosteron rekabetçi eğilimleri güçlendirebilmekte, dopamin ise trollük ve manipülasyon gibi davranışlarda ödül mekanizmasını harekete geçirmektedir. Hormonal dengeler ve dopamin döngüsü ise trollükten alınan hazı pekiştirerek itibar kaybını “zamansal indirgeme” ile görünmez kılmaktadır (Mazur & Booth, 1998; Buckels vd., 2014; Mazas vd., 2000; Blair vd., 2020). Dijital “sürüdeki” akran varlığı ise ventral striatumu tetikleyerek risk almayı bir sosyal ödüle dönüştürürken dijital platformların toksik yapısı savaşçı gen olarak bilinen Monoamin Oksidaz A (MAOA) geni gibi genetik faktörlerin çevresel koşullarla etkileşim içinde saldırgan davranışlara yatkınlığı etkilediği belirtilmektedir (Caspi vd., 2002; Burt, 2022).

## 2.2. Sosyolojik ve Dijital Yapısal Etkiler

Dijital antisosyal eylemler şahsi eğilimlerden ziyade, teknolojik imkanların maliyet dengesini bozduğu bir zeminde yeşermektedir. Durkheim’ın (1893)

orijinal tarihli eserinde klasik anomi kavramı, normların belirsizleştiği geçiş dönemlerinde saldırganlığın arttığını öngörmektedir (Durkheim, 2023). Gelfand vd. (2024) bu durumu “kültürel evrimsel uyumsuzluk” olarak tanımlar. Sosyal ikilemlerde görüldüğü gibi, norm boşluğundaki aktörler kendi bencil dürtülerini genele yansıtarak eylemlerini meşrulaştırırlar bu durum dijital ortamdaki saldırganlığın bir norm yanığıyla rasyonelleştirilmesine zemin hazırlar. (Briscese vd., 2024). Eylem artık meşrulaşmaktadır. Sosyal ipuçlarının dijital platformlarda olmaması otosansürü ve sosyal maliyeti minimize ederken (Suler, 2004), “dijital dağınmık kolektivite” veya “dijital kalabalık sayesinde” bireyin sorumluluğu seyreltilerek eylemi düşük duygusal maliyetli kolektif bir hazza dönüştürür (Hossain vd., 2025). Belirsizliğin sunduğu “zihinsel hareket alanı” ise antisoyal tercihleri stratejik bir kazanıma dönüştürmektedir (Olschewski vd., 2019).

Sosyal medya algoritmaları ise bu denklemin fayda kısmını yöneterek süreci kışkırtmaktadır. Bu platformlarda kullanılan algoritmalar etkileşimi ödüllendirerek provokatif ve kutuplaştırıcı söylemleri öne çıkarır ve maruz kalınan yoğun görünürlük baskısı aktörleri sistemle oynamaya iterek dijital kamusal alanı bir kalitesiz ve saldırgan içeriğin baskın olduğu “limon piyasasına” dönüştürür (AlMuhanna vd. 2023; Ren vd., 2024; Russ-Mohl, 2015). Kitle içindeki sorumsuzluk eğilimi ve dışlanma korkusu, bireyleri kendi fikirleri olmasa dahi yangın fırtınalarının rasyonel birer parçası haline getirmektedir (Le Bon, 2024; Wilczek, 2016). “Kırık camlar” teorisine ait bulgularla paralel olarak desteklendiği gibi, zehirli bir tartışma bağlamı, en sıradan kullanıcıyı bile bir trolle dönüştürebilir (Cheng vd., 2017). Nihai olarak, iktidar ve disiplin kavramları, dijital linç ve iptal kültürünü, modern bir denetim mekanizmasına dönüştürür (Foucault, 2012). Bu noktada bireyler, altruistik cezalandırma motivasyonu ile hiçbir maddi çıkar gözetmeksizin normları korumayı ahlaki bir ödev olarak üstlenirken troller tam da bu “samimi ahlaki endişeyi” (sincere concern) kendileri için bir alay ve istismar malzemesi olarak kullanarak sistemi sabote ederler (Fehr & Gächter, 2002; Kaplan, 2021).

### 3. Kültürel Boyutlar ve Dijital Davranışlar

Çevrim içi antisoyal davranışların bireylerin içinde buldukları kültürel değer sistemlerinden bağımsız düşünülemez. Gelfand vd. (2011) ve Hofstede (2001) tarafından sunulan boyutlar, davranışların hangi durumlarda tolere edileceğini belirleyen temel çerçevelerdir. Rekabetçi teşviklerin yardımlaşmayı azalttığı, aktörlerin yetkinlik odaklı itibar yönetimi için sosyal bağları koparmayı rasyonel gördüğünü bilinmektedir (Riyanto & Yeo, 2025) Bu bağlamda, dijital ortamlardaki eylem maliyetleri evrensel bir matematikten ziyade, kültürel kodlara gömülü bağlamsal algıya dayanmaktadır (Levinson & Peng,

2006). Örneğin bireyci kültürlerde ifade özgürlüğü ve kişisel özerkliğin ön planda olması, ghosting ve trollük gibi davranışların “bireysel tercih” olarak çerçevelenip daha düşük sosyal maliyetle gerçekleştirilmesine zemin hazırlamaktadır (Hofstede, 2001). Bu tür bireyci ortamlarda rekabetçi teşvik yapıları yardımlaşma eğilimini azaltırken (Riyanto & Yeo, 2025); aktörler “yetkinlik ve güç” sinyali veren bir itibar yönetimi uğruna sosyal bağları koparmayı ve antisosyal eylemlerde bulunmayı rasyonel görebilmektedir (Krueger vd., 2020). Aynı zamanda bu yapılarda manipülasyon, bireylerin dijital belirsizlik altında kendi bencil dürtülerini genele yansıttıkları ve normalleştirdikleri pragmatik bir mekanizmayla şekillenmektedir (Briscese vd., 2024).

Buna karşılık kolektivist ve norm sıklığı yüksek ortamlarda dinamikler tamamen farklılaşmaktadır. Bu kültürlerde dijital linç veya iptal kültürü, bireysel bir sapmadan ziyade; grubun kimliğine ait “davranışsal reçeteleri” (Akerlof & Kranton, 2000) ve toplumsal düzeni korumayı amaçlayan meşru bir “altruistik cezalandırma” aracı olarak rasyonelleştirilmektedir (Fehr & Gächter, 2002; Gelfand vd., 2024; Singelis vd., 1995)

Güç mesafesi boyutu da toplumlarda eylemlerin meşruiyetini belirleyen sessiz bir kültürel mimardır. Hiyerarşinin katı olduğu ortamlarda manipülatif kontrol rasyonel bir tercihe dönüşmekte; özellikle yoğun “meşruiyet baskısı” altında etik dışı davranışlar kolayca normalleşebilmektedir (Hofstede, 2001; Ren vd., 2024). Benzer biçimde, asimetrik güç yapılarında toplumsal refahı düşüren rüşvet maksimizasyonu ve sömürü stratejileri daha tolere edilebilir hale gelmektedir (Balafoutas vd., 2021). Öte yandan dijital troller, hiyerarşiden bağımsız olarak, bireylerin veya kurumların sergilediği ciddiyeti ve “samimi endişeyi” (sincere concern) doğrudan bir alay ve istismar hedefine dönüştürmektedir (Kaplan, 2021).

Sıkı normlara sahip toplumlarda açık saldırganlık hızla sorgulanarak yüksek sosyal maliyetle karşılanmakta ve kolektif dışlama paradoksal bir meşruiyet kazanmaktadır (Gelfand vd., 2011). Bu yapılarda üçüncü taraf cezalandırması bu yapılarda yaygındır (Gelfand vd., 2024). Bireyler, düzeni korumayı bir ödev addederek altruistik cezalandırma saikiyle linç kampanyalarına eklenir (Fehr ve Gächter, 2002).

Dijital platformların tasarımından kaynaklanan denetim eksikliği ve zayıf sosyal bağlar, tıpkı fiziksel mekanlardaki gözetim zafiyetleri gibi, manipülatif eylemlerin maliyetini düşürerek antisosyal davranışlara yönelik tolerans eşiğini yükseltmektedir (Moqadam & Nubani, 2022). Güç, rekabet ve başarının yüceltildiği maskülen kültürlerde çevrim içi linç ve doxxing, grubun üstünlüğünü pekiştiren meşru birer cezalandırma aracı olarak işlev görmektedir (Hofstede,

2001; Loveluck, 2020). Bu rekabetçi teşvik yapıları altında dijital etkileşim, yardımlaşmadan ziyade sıfır toplamlı bir oyuna dönüşmekte; saldırganlık ise statü kazanımı için göze alınan rasyonel bir yatırım maliyeti halini almaktadır (Riyanto & Yeo, 2025).

Buna karşın, empati ve uyumun ön planda olduğu feminen normların baskın olduğu toplumlarda bu tür davranışlar, gündelik ahlakın ağır bir ihlali olarak kodlanmakta ve uyuşmazlıklar çatışmadan ziyade müzakereyle çözülmektedir (Adams & Millie, 2021; Goldstein, 2003; Hofstede, 2001). Dolayısıyla, rekabetçi ortamlarda trollük bir üstünlük ve dikkat çekme aracı olarak ödüllendirilirken; uyum odaklı bağlamlarda bu benmerkezcî dürtüler, sürdürülebilir olmayan kibirli bir gurur olarak etiketlenip sosyal dışlanmayla sonuçlanmaktadır (Buckels vd., 2014; Costello vd., 2019; Hardaker, 2010).

#### 4. Dijital Antisosyal Davranış Türlerinin Karşılaştırmalı Analizi

Dijital antisosyal davranışlar, dışarıdan irrasyonel görünse de bireyin fayda ve maliyet ekseninde şekillenen tutarlı birer stratejik tercihtir. Bu kapsamda farklı antisosyal davranışların analiz edilmesi büyük önem arz etmektedir. Örneğin nefret söylemi, dış grubu ötekileştirerek iç grubun ‘kimlik faydasını’ pekiştirme girişimidir (Akerlof & Kranton, 2000; Tajfel vd., 2001). Ghosting, ilişkisel maliyetlerden kaçınmayı hedefleyen tek taraflı bir ‘Parayı Al ve Kaç’ stratejisidir (Ibáñez vd., 2016). Çevrim içi linç, ‘altruistik cezalandırma’ ritüeli üzerinden katılımcıya ahlaki üstünlük ve grup onayı sağlayan bir sürü davranıştır (Fehr & Gächter, 2002; Wilczek, 2016). Manipülasyon ve gaslighting ise, belirsizliğin sunduğu ‘zihinsel hareket alanını’ kullanarak kurbanın bilişsel kontrolünü ele geçirmeye odaklanır (Olschewski vd., 2019; Sweet, 2019).

Öte yandan siber zorbalık, acımasız bir yetkinlik ve statü inşası iken (Krueger vd., 2020); trollük gündelik sadizmin dijital bir yansımasıdır (Buckels vd., 2014). Bu provokatif eylem, hedefteki ciddiyeti ve ‘samimi endişeyi’ manipüle ederek ‘alay edilebilir bir tepki’ üretmeyi amaçlar (Kaplan, 2021). Nihayetinde bu pratikler, eylemi gerçekleştiren bireyi aşarak anonim bir kalabalıkla paylaşılan ‘kolektif özgün bir haz’ mekanizması yaratmaktadır (Hossain vd., 2025).

### 5. Kültürlerarası Karşılaştırmalı Bulgular

#### 5.1. Benzer Davranışlar, Farklı Anlamlar

Dijital eylemlerin maliyet ve fayda ağırlıkları evrensel bir matematikle değil, kültürel ve bağlamsal ipuçlarıyla hesaplanmaktadır (Levinson & Peng, 2006). Örneğin, ghosting eylemi bireyci yapılarda düşük maliyetli bir kişisel sınır koruma stratejisi olarak görülür. Buna karşın, kolektivist bağlamlarda

aynı eylem, ilişkisel sermayenin ağır bir ihlali olarak kodlanmakta ve sosyal maliyeti artmaktadır (Hofstede, 2001; Park vd., 2021).

Benzer şekilde, norm sıklığı yüksek toplumlarda çevrim içi linç, düzeni restore eden meşru bir sosyal denetim mekanizması işlevi görmektedir (Gelfand vd., 2024). Bu toplumlarda birey, norm ihlallerine yönelik saldırganlığı, toplumu korumayı amaçlayan özgecil bir cezalandırma (altruistic punishment) hazzıyla rasyonelleştirir (Fehr & Gächter, 2002). Bireysel hakların ön planda olduğu esnek kültürlerde ise bu pratik, faile itibar riski yükleyen orantısız bir dijital şiddet olarak yorumlanmaktadır (Traversa, 2022). Neticede bu tür kolektif saldırılar, rasyoneliteyi bireysel bir fayda-maliyet hesabından çıkararak; anonim dijital kalabalıkların paylaştığı ‘kolektif özgün bir haz’ zeminine taşımaktadır (Hossain vd., 2025).

## 5.2. Meşruiyet Farklarının Kültürel Kaynakları

Antisosyal eylemlerin meşruiyet zemini, norm sıklığı ve hiyerarşik bariyerlerin kesişim kümesinde inşa edilmektedir. Toplumsal tehdit algısı ile normların ‘sıklık’ derecesi arasındaki doğrudan bağ, ihlallere verilen tepkinin şiddetini belirlemektedir (Gelfand vd., 2024). Bireyin güvenliğini doğrudan ve açıkça tehdit eden doxxing gibi eylemler yüksek sosyal maliyet üretirken; manipülasyon ve gaslighting gibi örtük davranışlarda fayda-maliyet dengesi değişmektedir. Dijital belirsizliğin faile sunduğu ‘zihinsel hareket alanı’, niyetleri gizleme konforu sağlayarak sosyal yaptırım maliyetlerinden kaçınmayı kolaylaştırmaktadır (Olschewski vd., 2019).

Özellikle katı hiyerarşilerde bu tür stratejik ve manipülatif davranışlar çok daha kolay normalleşebilmektedir (Christie & Geis, 2013; Hofstede, 2001). Yoğun meşruiyet baskısı altındaki aktörler, sistemde kalabilmek adına aldatıcı hamleleri rasyonel birer zorunluluk olarak görmektedir (Ren vd., 2024). Güç asimetrisinin baskın olduğu bu toplumlarda, toplumsal refahı feda edip kişisel çıkarı önceleyen sömürücü eylemler adeta yazılmamış kurallara dönüşmektedir (Balafoutas vd., 2021). Dolayısıyla asıl mesele sadece davranışları sınıflandırmak değil; dijital ortamın zayıf ‘okunabilirliği’ ve denetim boşluğunun bu eylemleri nasıl mümkün kıldığını anlamaktır (Moqadam & Nubani, 2022). Bu bağlamda stratejik müdahaleler, bireylerin fayda ve maliyet algısını manipüle eden bu yapısal ve mimari kusurlara odaklanmalıdır.

## 6. Kültürel Kodlar ve Tercih Fonksiyonları

Davranışsal mikroiktisadi yaklaşım, tercih fonksiyonlarını salt klasik değişkenlerden ziyade kültürel beklentileri içerecek şekilde yeniden tanımlamakta; kararların evrensel bir akıldan ziyade sosyal olarak gömülü

ipuçlarıyla şekillendiğini göstermektedir (Levinson & Peng, 2006). Norm etkileşimlerinin fayda çerçevesini çizdiği bu yapıda, dijital davranışlar aslında kültürel kodlarla yazılan dinamik bir hesaplama sürecidir (Becker, 1996; Gelfand vd., 2024; North, 1990).

Bu açıdan bireycilik-kolektivizm eksenini, antisosyal eylemlerden elde edilen faydanın niteliğini doğrudan belirlemektedir. Bireyci kültürlerde dijital ortamlar birer kişisel görünürlük alanı olarak algılanmakta; rekabetçi teşvik yapıları yardımlaşmayı sönmümlendirerek, saldırgan davranışları faile 'yetkinlik' ve güç sinyali veren rasyonel birer getiriye dönüştürmektedir (Krueger vd., 2020; Riyanto & Yeo, 2025). Buna karşın kolektivist bağlamlarda, grup uyumunu bozan hamleler son derece yüksek sosyal maliyetler üretir (Hofstede, 2001; Singelis vd., 1995). Başkalarının refahını gözetmenin bireyin kendi mutluluğunu da artırdığı göz önüne alındığında, uyum odaklı bu kültürlerde antisosyal bir eylem, doğrudan içsel tatmin kaybına dönüşerek bireyin tercih setini daraltmaktadır (Iwasaki, 2023).

Güç mesafesi, dijital mecralarda eylemlerin meşruiyetini belirleyen sessiz bir süzgeçtir. Otoritenin sorgulanmadığı hiyerarşik yapılarda manipülatif eylemler, artan görünürlük ve meşruiyet baskısı altında hayatta kalabilmek için rasyonel birer stratejiye dönüşmektedir (Hofstede, 2001; Ren vd., 2024). Aynı zamanda bu ihlaller, grubun davranışsal reçetelerini ve sınırlarını koruyan bir 'kimlik ekonomisi' aracı olarak da işlev görebilmektedir (Akerlof & Kranton, 2000). Düşük güç mesafesine sahip kültürlerde ise bu edimler doğrudan hak ihlali olarak kodlanmakta ve bireye dramatik maliyetler yüklemektedir (Hofstede, 2001; Schwartz, 2012).

Norm sıklığı, yaptırımları öngörülebilir kılarak yakalanma riskini artırmaktadır (Gelfand vd., 2011; Gelfand vd., 2024). Buna karşın dijitalin esnek ve belirsiz dokusu, bireylerin kendi bencil dürtülerini genele yansıtmasına (projection) zemin hazırlamaktadır (Briscese vd., 2024). Beklenti teorisindeki standart kayıptan kaçınma (loss aversion) reflekslerinin aksine; risk algısındaki bu bilişsel sapma ve çerçeveleme farklılıkları, nefret söylemini ve saldırganlığı fail için 'düşük maliyetli' bir seçeneğe dönüştürmektedir (Costello vd., 2019).

Sonuç olarak, dijital davranışların mikroiktisadi rasyonalitesi evrensel olarak sabit değildir; bu değer, soyut matematikten ziyade yerel ve kültürel ipuçlarıyla kurgulanmaktadır (Levinson & Peng, 2006). Örneğin turnuva tipi rekabetçi teşvik yapılarında saldırganlık bir tür 'performans' olarak ödüllendirilirken, işbirlikçi iklimlerde sert yaptırımlarla karşılaşmaktadır (Riyanto & Yeo, 2025). Neticede kültür, sadece bireysel tercih fonksiyonlarını biçimlendirmekle kalmaz; dijital platform kaynaklı negatif dışsallıkların şiddetini ve toplumsal refahın sınırlarını da doğrudan belirler.

## 7. Genel Değerlendirme ve Politika Çıkarımları

Bu kitap bölümü, aynı dijital antisosyal davranışların, farklı kültürel bağlamlarda farklı marjinal fayda ve maliyetler ürettiğini ve dolayısıyla davranışın mikroiktisadi anlamda rasyonelliğin kültürel normlara göre değişkenlik gösterdiğini ortaya koymaktadır. Bulgular, davranışların salt bireysel patolojiden ziyade, eski zihinsel donanımların dijital kuralızsızlıkla girdiği kültürel evrimsel uyumsuzluk kavramının bir ürünü olduğuna işaret etmektedir (Gelfand vd., 2024). Antisosyal davranış, kültürel tercih fonksiyonları ile dijital dağınık kolektivite yapısının kesişiminde şekillenmektedir (Hossain vd., 2025). Mikroiktisadi açıdan fayda maksimizasyonu evrensel değildir. Kültürel normlar ve güç hiyerarşileriyle sürekli yeniden kodlanan bir süreçtir (Levinson & Peng, 2006; Becker, 1996; North, 1990). Dijital platformlar, anonimlik yoluyla sosyal maliyeti düşürürken, yetkinlik sinyali gibi algılanan getirileri ise maksimize etmektedir (Krueger vd., 2020). Neticede meşruiyet baskısı, belirli bağlamlarda antisosyal stratejileri rasyonel birer hamleye dönüştürmektedir (Ren vd., 2024).

Antisosyal kişilik eğilimleri, bireylerin tercih fonksiyonlarında ‘zamansal indirgeme dürtüsellliği’ (temporal discounting impulsivity) yaratarak, onları uzun vadeli maliyetleri hesaba katmadan anlık hazlara yöneltilmektedir (White vd., 2013). Özellikle dijital platformlardaki negatif tartışma bağlamı saldırganlığı daha da tetiklemektedir (Cheng vd., 2017). Bu durum, kimliksizleşme ve anonimliğin de etkisiyle eylemlerin maliyet hesabını zorlaştırmakta ve bireyi rasyonel olmayan bir sürü psikolojisine ve kolektif bulaşmaya itmektedir (Hossain vd., 2025; Wilczek, 2016). Nitekim risk algısındaki çerçeveleme etkilerinde de görüldüğü üzere, bu eylemler salt bireysel bir patoloji değildir; aksine, zihnin dijital alanın değişen kurallarına uyum sağlamaya çalışırken ürettiği bağlama duyarlı rasyonalite biçimleridir (Costello vd., 2019).

Bu çerçevede, dijital antisosyal davranışlarla mücadeleyle yönelik politika çıkarımları, sadece bireyleri cezalandırmanın ötesine geçmelidir. Kök nedenleri çözmeden sadece hatalı bireyleri düzeltmeye çalışan kısıtlı müdahaleler, maliyet-etkinlik analizlerine göre ekonomik olarak oldukça verimsizdir (Foster, 2010). Bunun yerine politikalar, çevresel teşvik yapılarını ve bireylerin tercih fonksiyonlarını dönüştürmeye odaklanmalı; prososyal davranışların bireysel ‘iyi oluş’ üzerindeki pozitif etkisini vurgulayarak dijital dengeyi yeniden iyilik lehine kurmalıdır (Iwasaki, 2023).

### 7.1. Önleme: Kültüre Duyarlı Tercih Yapılarının Yeniden Tasarımı

Dijital antisosyal eylemlerle mücadelede ‘herkese uyan tek çözüm’ paradigması işlevsizdir. Kültürel farklılıklar, bireylerin ekonomik değerlendirme

süreçlerini ve fayda-maliyet algılarını doğrudan değiştirmektedir; bu nedenle önleyici müdahalelerin mutlaka yerleştirilmesi gerekmektedir (Gelfand vd., 2011; Hofstede, 2001; Levinson & Peng, 2006).

Politika yapıcılar sadece sonuçları cezalandırmak yerine, davranışı teşvik eden kök nedenlere odaklanmalıdır; aksi takdirde müdahaleler ekonomik olarak verimsiz kalacaktır (Foster, 2010). Antisozyal aktörlerin karar alma süreçlerinde belirgin bir ‘anlık ödül yanlılığı’ (immediate reward bias) bulunmaktadır. Bu durum göz önüne alındığında, gecikmeli cezalar yerine anında devreye giren caydırıcı geri bildirim mekanizmaları kurmak hayati önem taşımaktadır (Mazas vd., 2000).

Platform düzeyinde ise sosyal medyanın yarattığı ‘dağınık kolektivitenin’ kontrol altına alınması gerekmektedir. Bu tür topluluk temelli müdahaleler, anonim kalabalıkların antisozyal eylemlerden elde ettiği ‘kolektif özgün hazzı’ seyretecektir (Hossain vd., 2025). Eş zamanlı olarak mağduriyetin görünür kılınması, faillerdeki ‘herkes yapıyor’ yanlılığını kırarak topluluk normlarının doğru yönde yeniden kalibre edilmesini sağlayacaktır (Briscese vd., 2024).

Son olarak, müdahalelerin başarısında kültürel uyum kritik bir değişkendir. Norm sıklığının yüksek olduğu toplumlarda, grup kimliklerini tehdit eden ani ve sert yaptırımlar kültürel ataleti tetikleyerek değişimi baltalayabilir; bu nedenle kolektif yaptırımlarda ölçülü olunmalıdır (Gelfand vd., 2024). Esnek yapılarda ise durum farklıdır. Bu ortamlarda negatif dışsallıklar açıkça görünür kılınmalı ve prososyal davranışların bireyin kendi ‘iyi oluşu’ üzerindeki pozitif etkisi vurgulanarak dijital teşvikler yeniden tasarlanmalıdır (Iwasaki, 2023).

## **7.2. Platform Sorumluluğu: Teşvik Yapıları ve Algoritmik Tercihler**

Sosyal medya mecraları, bireylerin tercih fonksiyonlarını doğrudan şekillendiren güçlü teşvik merkezleridir. Görünürlük odaklı algoritmalar, bilginin yayılımına dayalı ‘mevcudiyet kaskadları’ yaratarak saldırganlığı ödüllendirmekte ve antisozyal davranışın marjinal faydasını yapay şekilde artırmaktadır (Wilczek, 2016). Bu teşvik yapısı, çevrim içi kimlik kaybının etkisiyle birleştiğinde, yıkıcı eylemler birey için iktisadi olarak rasyonel bir tercihe dönüşmektedir (Suler, 2004).

Platformların sağladığı ‘dijital dağınık kolektivite’ ise bireysel sorumluluğu büyük ölçüde seyretektir. Eş zamanlı olarak anonimlik eylemin sosyal maliyetini düşürürken, failerin anonim kalabalıktan elde ettiği ‘kolektif özgün hazzı’ maksimize etmektedir (Hossain vd., 2025). Bu nedenle platform politikaları salt içerik temizliğine değil, bozulan bu maliyet-getiri dengesini onarmaya odaklanmalıdır. Antisozyal aktörlerin karar alma süreçlerindeki belirgin ‘gelecek miyopluğu’, uzun vadeli yaptırım tehditlerini caydırıcılıktan

uzak ve anlamsız kılmaktadır (Mazas vd., 2000; Petry, 2002). Çözüm, dijital ortamı toksik içeriklerin baskın olduğu bir ‘limon piyasası’ olmaktan çıkaracak anlık geri bildirim ve şeffaf hesap verebilirlik mekanizmalarıdır (Russ-Mohl, 2015). Kullanıcılar üzerinde salt bir meşruiyet baskısı kurmak yerine, bağlama duyarlı tasarımlarla ihlal fırsatlarını daraltmak, prososyal iş birliğini çok daha rasyonel bir seçeneğe dönüştürecektir (Ren vd., 2024).

Ancak platformların bu müdahalelerinde ‘herkese uyan tek tip’ ayarlamalardan kaçınılmalıdır. Bireylerin risk, maliyet ve faydaya değer biçme süreçleri evrensel değildir; bu nedenle moderasyon politikaları mutlaka yerel ve kültürel farklılıkları dikkate alınmalıdır (Levinson & Peng, 2006). Aksi takdirde, yerel dinamikleri ve grup aidiyetlerini tehdit eden tepeden inme müdahaleler, güçlü bir ‘kültürel atalet’ (cultural inertia) yaratarak değişimi baltalayabilir ve ‘kültürel evrimsel uyumsuzluk’ (mismatch) krizlerini daha da derinleştirebilir (Gelfand vd., 2024).

### 7.3. Dijital Okuryazarlık: Uzun Vadeli Tercih Dönüşümü

Bir diğer politika önerisi olarak dijital okuryazarlık, salt teknik bir donanımdan ziyade, bireysel tercih fonksiyonlarını uzun vadede dönüştürecek stratejik bir ‘kültürel yetkinlik’ aracı olarak kurgulanmalıdır (Levinson & Peng, 2006). Sosyal oyun teorisi perspektifinden bakıldığında, dijital platformların mevcut yapısı ahlaki boyutu gölgede bırakarak çoğunlukla rekabetçi yetkinliği ödüllendirmektedir (Krueger vd., 2020). Bu nedenle müdahaleler, karşılıklılık ilkesini (Altın Kural) dijital etkileşime entegre ederek; aktörleri miyopik ‘lulz’ arayışından uzaklaştırıp, kendi iyi oluşlarını önceleyen kararlara yönlendirmelidir (Adams & Millie, 2021; Buckels vd., 2014). Belirsizlik ortamında bireylerin kendi bencil tercihlerini genele yansıtma (projection) eğilimi, doğru norm bilgisinin görünür kılınmasıyla yeniden kalibre edilebilmektedir (Briscese vd., 2024). Nihayetinde, başkasının refahını da içeren prososyal tercihlerin  $WB(L, a_i)$  bireysel mutluluğu doğrudan artırdığı gerçeği merkeze alınmalı (Iwasaki, 2023), bu tür kapsayıcı değerlemeler üzerinden dijital ekosistemde sürdürülebilir bir sosyal sermaye inşa edilmelidir (Gelfand vd., 2024).

### Sonuç

Dijital antisosyal davranışlar (trollükten linç pratiklerine) basit birer psikolojik sapma değildir, kültürel kodlar ve teknolojik teşviklerle harmanlanan stratejik tercih fonksiyonlarının çıktılarıdır. Bu analizin ortaya koyduğu en kritik gerçek, insan zihninin evrimsel mirası ile dijital çağın dinamikleri arasında yaşanan ‘kültürel evrimsel uyumsuzluktur’ (Gelfand vd., 2024). Bireyin mikro dünyasında rasyonel bir mantığa dayanan bu yıkıcı eylemler; anonimliğin düşürdüğü sosyal maliyetler ile statü ve grup aidiyeti gibi maksimize edilen

faydaların kesişiminde hayat bulmaktadır. Yüz yüze etkileşimin ahlaki fren mekanizmaları bu belirsiz sahada işlevsizleşirken, gelecek miyopluğu yaşayan aktörler kısa vadeli anlık ödüller uğruna uzun vadeli sosyal sermayelerini feda etmektedir.

Bu nedenle çözüm, bozulan tercih fonksiyonlarını onaracak üçlü bir mimariyi zorunlu kılmaktadır. Birincisi, dijital düzenlemelerin kültürel boyutlar merkeze alınarak, yerel norm sıklığına ve kültürel hassasiyetlere duyarlı hale getirilmesidir (Gelfand vd., 2011; Hofstede, 2001). İkinci olarak, dijital platformların pasif aracı konumundan çıkıp aktif birer politika aktörü olarak sorumluluk üstlenmesi gerekmektedir. Saldırganlığı ve viraliteyi ödüllendiren algoritmalar dijital alanı kalitesiz bir 'limon piyasasına' dönüştürdüğünden; platformların salt etkileşimden ziyade empatiyi ve iş birliğini ödüllendiren teşvik yapıları inşa etmeleri şarttır (Ren vd., 2024; Russ-Mohl, 2015; Wilczek, 2016). Üçüncüsü, eğitim politikaları, prososyal davranışların bireysel iyi oluşu artırdığı gerçeğini merkeze alarak algılanan faydayı dönüştürmeli ve dijital belirsizliğin yarattığı norm yanılgılarını yeniden kalibre etmelidir (Briscese vd., 2024; Iwasaki, 2023).

Nihayetinde bu bütüncül yaklaşım, sadece dijital riskleri azaltmayı hedefleyen bir savunma stratejisi değildir. Bu dijital alanın 'vahşi batı' kurallarından arındırılıp, güvenin, hesap verebilirliğin ve karşılıklılığın yeniden tesis edildiği, daha sürdürülebilir, kapsayıcı ve insani bir dijital kamusal alanın inşası için bir çağrıdır.

## Kaynakça

- Adams, J., & Millie, A. (2021). Everyday moral judgements of anti-social behaviour. *Crime Prevention and Community Safety*, 23, 56-68. <https://doi.org/10.1057/s41300-020-00106-6>
- Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The quarterly journal of economics*, 115(3), 715-753. <https://doi.org/10.1162/003355300554881>
- AlMuhanna, N., Hall, W., & Millard, D. E. (2023). Fear of the dark: a cross-cultural study into how perceptions of antisocial behaviour impact the acceptance and use of Twitter. *Behaviour & Information Technology*, 42(8), 1180-1193. <https://doi.org/10.1080/0144929X.2022.2064766>
- Balafoutas, L., García-Gallego, A., Georgantzis, N., Jaber-Lopez, T., & Mitrokostas, E. (2021). Psychopathy and economic behavior among prison inmates: An experiment. *Frontiers in Psychology*, 12, 732184. <https://doi.org/10.3389/fpsyg.2021.732184>
- Becker, L. C. (1996). Trust as noncognitive security about motives. *Ethics*, 107(1), 43-61. <https://doi.org/10.1086/233696>
- Blair, R. J. R., Bashford-Largo, J., Zhang, R., Lukoff, J., Elowsky, J. S., Leibenluft, E., ... & Blair, K. S. (2020). Temporal discounting impulsivity and its association with conduct disorder and irritability. *Journal of Child and Adolescent Psychopharmacology*, 30(9), 542-548. <https://doi.org/10.1089/cap.2020.0001>
- Bowlby, J., Ainsworth, M., & Bretherton, I. (1992). The origins of attachment theory: John Bowlby and Mary Ainsworth. *Developmental Psychology*, 28(5), 759-775.
- Briscese, G., Grignani, M., Lacetera, N., Macis, M., & Tonin, M. (2024). Norm misperceptions in social dilemmas: the role of preferences, heuristics and experiences. *Journal of Behavioral Economics for Policy*, Vol.8, Special Issue1, 9-24.
- Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences*, 67, 97-102. <https://doi.org/10.1016/j.paid.2014.01.016>
- Burt, S. A. (2022). The genetic, environmental, and cultural forces influencing youth antisocial behavior are tightly intertwined. *Annual Review of Clinical Psychology*, 18, 155-178. <https://doi.org/10.1146/annurev-clinpsy-072220-015507>
- Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., Taylor, A. & Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, 297(5582), 851-854. <https://doi.org/10.1126/science.1072290>
- Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2017). Anyone can become a troll: Causes of trolling behavior in online dis-

- cussions. *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW '17)*, 1217–1230. <https://doi.org/10.1145/2998181.2998213>
- Christie, R., & Geis, F. L. (2013). *Studies in machiavellianism*. Academic Press, 2013.
- Costello, T. H., Smith, S. F., Bowes, S. M., Riley, S., Berns, G. S., & Lilienfeld, S. O. (2019). Risky business: Psychopathy, framing effects, and financial outcomes. *Journal of Research in Personality*, 78, 125–132. <https://doi.org/10.1016/j.jrp.2018.11.006>
- Durkheim, É. (1893). *De la division du travail*. FB Editions.
- Durkheim, E. (2023). The division of labour in society. In *Social theory re-wired* (pp. 15-34). Routledge.
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137-140. <https://doi.org/10.1038/415137a>
- Foucault, M. (2012). *Discipline and punish: The birth of the prison*. Vintage. Knopf Doubleday Publishing Group, 2012.
- Freud, S. (2015). *Civilization and its discontents*. Broadview Press.
- Foster, R. D. (2010). Resistance, justice, and commitment to change. *Human resource development quarterly*, 21(1), 3-39. <https://doi.org/10.1002/hrdq.20035>
- Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., ... & Yamaguchi, S. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, 332(6033), 1100-1104. <https://doi.org/10.1126/science.1197754>
- Gelfand, M. J., Gavrillets, S., & Nunn, N. (2024). Norm dynamics: Interdisciplinary perspectives on social norm emergence, persistence, and change. *Annual Review of Psychology*, 75, 341–378. <https://doi.org/10.1146/annurev-psych-033020-013319>
- Goldstein, A. P. (2003). *The psychology of group aggression*. John Wiley & Sons.
- Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of politeness research*, 6(2), 215-242. <http://dx.doi.org/10.1515/JPLR.2010.011>
- Hofstede, G. (2001). Culture's consequences: Comparing values, behaviors, institutions and organizations across nations. *International Educational and Professional*.
- Hossain, M. A., Quaddus, M., Akter, S., Mikalef, P., & Warren, M. (2025). Trolling in social media: A deindividuation and contagion perspective. *Information & Management*, 62, 104211. <https://doi.org/10.1016/j.im.2025.104211>
- Ibáñez, M. I., Sabater-Grande, G., Barreda-Tarrazona, I., Mezquita, L., López-Ovejero, S., Villa, H., ... & Georgantzis, N. (2016). Take the money and

- run: Psychopathic behavior in the trust game. *Frontiers in Psychology*, 7, 1866. <https://doi.org/10.3389/fpsyg.2016.01866>
- Iwasaki, M. (2023). Social preferences and well-being: theory and evidence. *Humanities and Social Sciences Communications*, 10, Article 318. <https://doi.org/10.1057/s41599-023-01782-z>
- Kaplan, J. (2021). Speech, mockery, and sincere concern: An account of trolling. *Public Affairs Quarterly*, 35(3), 204-227. <https://www.jstor.org/stable/10.2307/48628248>
- Keltner, D., Ellsworth, P. C., & Edwards, K. (1993). Beyond simple pessimism: Effects of sadness and anger on social perception. *Journal of Personality and Social Psychology*, 64(5), 740-752.
- Koenigs, M., Kruepke, M., & Newman, J. P. (2010). Economic decision-making in psychopathy: a comparison with ventromedial prefrontal lesion patients. *Neuropsychologia*, 48(7), 2198-2204.
- Krueger, J. I., Heck, P. R., Evans, A. M., & DiDonato, T. E. (2020). Social game theory: Preferences, perceptions, and choices. *European Review of Social Psychology*, 31(1), 222-253. <https://doi.org/10.1080/10463283.2020.1778249>
- Le Bon, G. (2024). The Crowd: A Study of the Popular Mind [1895]. *Early Media Effects Theory & the Suggestion Doctrine: Selected Readings, 1895-1935*, 13.
- Levinson, J. D., & Peng, K. (2006). Valuing cultural differences in behavioral economics. *The ICEAI Journal of Behavioral Finance*, 3(4), 32-47.
- Loveluck, B. (2020). The many shades of digital vigilantism. A typology of online self-justice. *Global crime*, 21(3-4), 213-241. <https://doi.org/10.1080/17440572.2019.1614444>
- Mazas, C. A., Finn, P. R., & Steinmetz, J. E. (2000). Decision-making biases, antisocial personality, and early-onset alcoholism. *Alcoholism: Clinical and Experimental Research*, 24(7), 1036-1040. <https://doi.org/10.1111/j.1530-0277.2000.tb04647.x>
- Mazur, A., & Booth, A. (1998). Testosterone and dominance in men. *Behavioral and brain sciences*, 21(3), 353-363. <https://doi.org/10.1017/S0140525X98001228>
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100(4), 674-701. <https://doi.org/10.1037/0033-295X.100.4.674>
- Moqadam, S., & Nubani, L. (2022). The impact of spatial changes of Shiraz's historic district on perceived anti-social behavior. *Sustainability*, 14(14), 8446. <https://doi.org/10.3390/su14148446>
- North, D. C. (1990). A transaction cost theory of politics. *Journal of theoretical politics*, 2(4), 355-367. <https://doi.org/10.1177/0951692890002004001>

- Olschewski, S., Dietsch, M., & Ludvig, E. A. (2019). Anti-social motives explain increased risk aversion for others in decisions from experience. *Judgment and Decision Making*, 14(1), 58–71.
- Park, M. A., Golden, K. J., Vizcaino-Vickers, S., Jidong, D., & Raj, S. (2021). Sociocultural values, attitudes and risk factors associated with adolescent cyberbullying in East Asia: A systematic review. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 15(1).<https://doi.org/10.5817/CP2021-1-5>
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature reviews neuroscience*, 9(2), 148-158. <https://doi.org/10.1038/nrn2317>
- Petry, N. M. (2002). A comparison of young, middle-aged, and older adult treatment-seeking pathological gamblers. *The gerontologist*, 42(1), 92-99. <https://doi.org/10.1093/geront/42.1.92>
- Ren, J., Wu, P., & Hou, L. (2024). Social media attention and corporate greenwashing: Evidence from China. *Corporate Social Responsibility and Environmental Management*, 31(6), 5446–5465. <https://doi.org/10.1002/csr.2875>
- Riyanto, Y. E., & Yeo, J. (2025). Cooperation or Competition? Social Identity and Incentives in Workgroups. Available at SSRN: <https://ssrn.com/abstract=5190980> or <http://dx.doi.org/10.2139/ssrn.5190980>
- Russ-Mohl, S. (2015). The (behavioral) economics of media accountability. *Facta Universitatis, Series: Philosophy, Sociology, Psychology and History*, 14(1), 15–27.
- Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., ... & Konty, M. (2012). Refining the theory of basic individual values. *Journal of personality and social psychology*, 103(4), 663. <https://doi.org/10.1037/a0029393>
- Seligman, M. E. (1975). *Helplessness* (pp. 166-188). San Francisco: Freeman.
- Shaver, P., & Hazan, C. (1987). Being lonely, falling in love. *Journal of Social Behavior and Personality*, 2(2), 105.
- Singelis, T. M., Triandis, H. C., Bhawuk, D. P., & Gelfand, M. J. (1995). Horizontal and vertical dimensions of individualism and collectivism: A theoretical and measurement refinement. *Cross-cultural research*, 29(3), 240-275. <https://doi.org/10.1177/106939719502900302>
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & behavior*, 7(3), 321-326. <https://doi.org/10.1089/1094931041291295>
- Sweet, P. L. (2019). The sociology of gaslighting. *American sociological review*, 84(5), 851-875. <https://doi.org/10.1177/0003122419874843>
- Tajfel, H., Turner, J., Austin, W. G., & Worchel, S. (2001). *An integrative theory of intergroup conflict*. Intergroup relations: Essential readings, 94-109.

- Traversa, M. (2022). *Is cancel culture validating? How collective validation mediates the relationship between cancel culture and collective action*. Simon Fraser University, Thesis M.A.
- White, S. E., Brislin, S. J., Meffert, H., Sinclair, S., & Blair, R. J. R. (2013). Callous-unemotional traits modulate the neural response associated with punishing another individual during social exchange: A preliminary investigation. *Journal of Personality Disorders*, 27(1), 99–112. <https://doi.org/10.1521/pedi.2013.27.1.99>
- Wilczek, B. (2016). Herd behaviour and path dependence in news markets: Towards an economic theory of scandal formation. *Journal of Interdisciplinary Economics*, 28(2), 137-167. <https://doi.org/10.1177/0260107916643469>
- Zuboff, S. (2019, January). Surveillance capitalism and the challenge of collective action. In *New labor forum* (Vol. 28, No. 1, pp. 10-29). Sage CA: Los Angeles, CA: Sage Publications. <https://doi.org/10.1177/1095796018819461>

